AD-A276 626

Application of Neural Networks to Seismic Signal Discrimination

James A. Cercone   V. Shane Foster
W. Mike Clark   Larry McCutchan
J. Joseph Fuller   John Martin
Stephen Goodman   G. Andy Cipriani
Don J. Smith   Heather Tanner
Barbara Crist

Tech Foundation, Inc.
Old Main
Montgomery, West Virginia 25136

DTIC
ELECTE
JAN 1 1 1994
S
A

15 May 1993

Scientific Report No. 1

94-01179

94 1 10 150

PHILLIPS LABORATORY
Directorate of Geophysics
AIR FORCE MATERIEL COMMAND
HANSCOM AIR FORCE BASE, MA 01731-3010

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Air Force or the U.S. Government.

This technical report has been reviewed and is approved for publication.
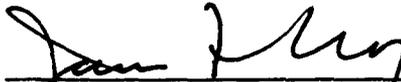
JAMES F. LEWKOWICZ
Contract Manager
Solid Earth Geophysics Branch
Earth Sciences Division
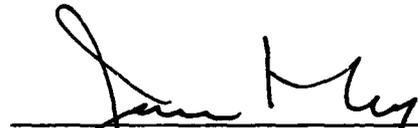
JAMES F. LEWKOWICZ
Branch Chief
Solid Earth Geophysics Branch
Earth Sciences Division

DONALD H. ECKHARDT, Director
Earth Sciences Division

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE 15 May 1993	3. REPORT TYPE AND DATES COVERED Scientific Report No. 1

4. TITLE AND SUBTITLE
Application of Neural Networks to Seismic Signal Discrimination

5. FUNDING NUMBERS
PE   61101E
PR 1D90 TA DA WU AA
Contract F19628-91-K-0039

6. AUTHOR(S)
James A. Cercone   Stephen Goodman   John Martin
W. Mike Clark   Don J. Smith   G. Andy Cipriani
J. Joseph Fuller   V. Shane Foster   Heather Tanner
                   Larry McCutchan   Barbara Crist

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Tech Foundation, Inc.
Old Main
Montgomery, West Virginia 25136

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Phillips Laboratory
29 Randolph Road
Hanscom AFB, MA 01731-3010

Contract Manager: James Lewkowicz/GPEH

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
PL-TR-93-2154

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This is the first Annual Technical Summary of the West Virginia Institute of Technology Applications of Neural Networks to Seismic Classification project. The first year of research focused on identification and collection of a suitable database, identification of parametric representation of the time series seismic waveforms, and the initial training and testing of neural networks for seismic event classification. It was necessary to utilize seismic events that had a high degree of reliability for accurate training of the neural networks. The seismic waveforms were obtained from the Center for Seismic Studies and were organized into three smaller databases for training and classification purposes. Unprocessed seismograms are not well suited for presentation to a neural network because of the large number of data points required to represent a seismic event in the time domain. Parametric representation of the seismic waveform numerically extracts those features of the waveform that enable accurate event classification. Sonograms and moment feature extraction are two of the several transformations investigated for parametric representation of a seismic event. This parametric representation of the seismic events provides adequate information.

14. SUBJECT TERMS
Neural networks   Data points
Signal discrimination   Parametric representation
Seismic events

15. NUMBER OF PAGES
76

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

TABLE OF CONTENTS

LIST OF TABLES

DTIC QUALITY INSPECTED 5

Accesion For		
NTIS CRA&I		☑
DTIC TAB		☐
Unannounced		☐
Justification		
By		
Distribution /		
Availability Codes		
Dist	Avail and / or Special	
A-1		

LIST OF FIGURES

EXECUTIVE SUMMARY

This is the first Annual Technical Summary of the West Virginia Institute of Technology Applications of Neural Networks to Seismic Classification project. The first year of research focused on identification and collection of a suitable database, identification of parametric representation of the time series seismic waveforms, and the initial training and testing of neural networks for seismic event classification. It was necessary to utilize seismic events that had a high degree of reliability for accurate training of the neural networks. The seismic waveforms were obtained from the Center for Seismic Studies and were organized into three smaller databases for training and classification purposes. Unprocessed seismograms are not well suited for presentation to a neural network because of the large number of data points required to represent a seismic event in the time domain. Parametric representation of the seismic waveform numerically extracts those features of the waveform that enable accurate event classification. Sonograms and moment feature extraction are two of the several transformations investigated for parametric representation of a seismic event. This parametric representation of the seismic events provides adequate information for accurate event classification, while significantly reducing the minimum size of the neural network. Preliminary results have achieved classification rate over 75% for the 5 class problem. Future work is focused on training and testing with larger datasets (300+ waveforms) and to determine the effects of seismic recording station location.

1.0   INTRODUCTION

Seismology is an old science with seismological records dating back to the BC era [5]. Detection and classification of seismic events have been extensively studied in recent years and require highly trained seismologists to accurately interpret seismic traces. The resurgence of neural network technology in the past decade has allowed re-examination of models and algorithms used for detection and classification of seismic traces. Neural networks provide a model free method of seismogram signal classification [10].

Figure 1 presents the seismic trace of a typical quarry blast. A seismologist, upon initial examination of the seismic trace, would probably suspect that this trace represents a quarry blast simply because of his training and experience. The seismologist would then try to identify the seismic origin by direct conformation with someone at the location of the blast site or from published schedules of such events. Unfortunately, there are thousands of seismic events daily, which makes it virtually impossible to identify each and every event with a high degree of certainty.

Without a-priori knowledge of the event classification, a seismologist would try to identify certain features of the seismic wave form. In particular, he would attempt to label certain phases of the wave form, the arrival time of the first surface waves, or P waves, the arrival time of secondary waves, S waves, and subsequent long waves, L and R. To a great extent, this is a subjective analysis at best. After phase identification, a tentative classification will be attached to the wave form. This process is often repeated by other seismologist for further verification.
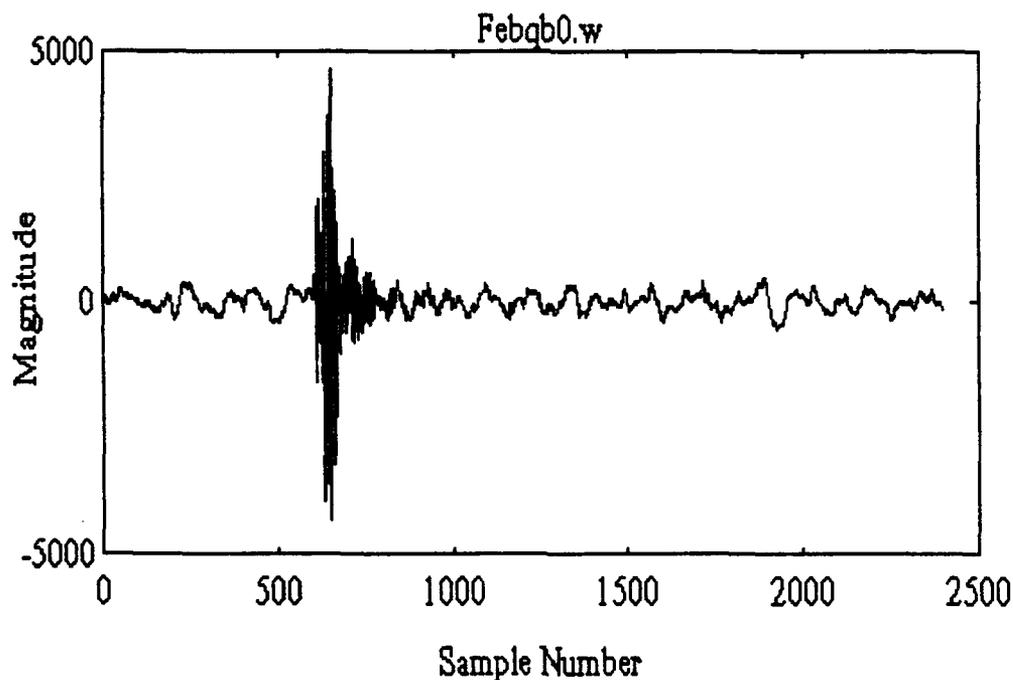
Figure 1 Quarry Blast

It has been necessary for the United State to operate various seismic surveillance systems over the course of the Cold War years in order to monitor nuclear treaty compliance. A seismic surveillance system is illustrated by the Intelligent Monitoring System (IMS) developed by ARPA for seismic data interpretation [4]. This system was designed to detect and locate seismic sources and help classify the event type. The IMS system has been integrated with neural network components by Lincoln Laboratory [2,3]. Figure 2 shows the functional elements of the modified IMS system developed.

Discrimination Review	Human Expert
Auto Event Discrimination	Simple Expert System/Pattern Classifier
Seismic Analyst Review	Human Expert
Auto Event List Processing	Expert Systemt
Signal Meas. & Phase ID	Simple Expert System
Detection Processing	Signal Processing
Seismic Signals	
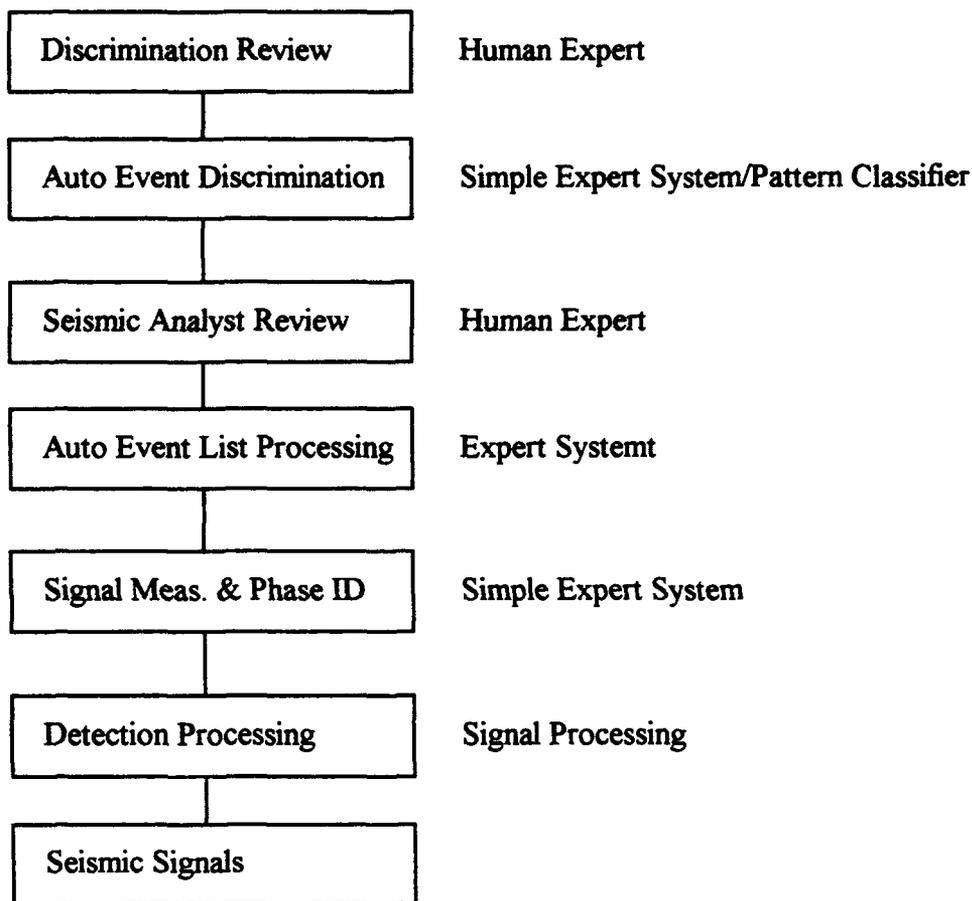
Figure 2 Modified Intelligent Monitoring System

This modified system adds expert system techniques not present in the original IMS. It should be noted that proper phase identification is still key to the modified system of Figure 2 developed at the Lincoln Labs.

Referring back to Figure 1, a non-seismologist could examine the seismic wave form and make several observations. These observations might include overall duration, rise time, the amount of clutter and possib;; the periodic components of the wave form.

These non-expert observations are subjective at best. To help understand what type of properties might serve as an indication of signal classification, an informal test was conducted in which 75 seismic traces composed of 5 events with 15 wave forms each, were shown to four professors and two graduate students. These wave forms were presented on a computer screen in flasn card style and in random order. After one training session, the group was able to correctly identify 65% to 90% of the presented wave forms with an average correct classification of about 75% (the author managed 65% recognition). Follow up discussion with the test group lead to some suggested heuristics that a non-seismologist and possibly a neural network might use for seismic event classification. These suggested heuristics are incorporated into the overall system design.

One of the key premises of this research was the deviation from the traditional methods used by seismologist to identify alternate methods that can implemented with neural networks.

1.2   Seismic Classification Problem

Seismic data analysis is roughly divided into two categories, exploration and non-deterministic event detection/classification. Starting with the activation of a seismic source, a received seismic trace s(t) can be considered the convolution of the original signal wavelet w(t) with a series of reflection coefficients, r(t)—earth's impulse response—with additive noise, expressed symbolically as:

$$s(t) = w(t) * r(t) + n(t).$$

If the earth's reflection coefficient's r(t) can be modeled, it is theoretically a simple matter of de-convolution to recover the original source w(t). Given w(t) as determined by at least three detectors, the source location and magnitude can be found. Exploration seismology involves processing the seismic trace to recover the reflection coefficients, r(t), with the original wavelet w(t) known. The reflection coefficients are then geologically interpreted to indicate the possible presence of natural resources such as gas and oil. Signal processing techniques aid in the determination of the reflection coefficients ranging from a simple one dimensional correlator to the incorporation of multidimensional fan filters for spatially different detectors [19].

Non-deterministic seismic event detection and subsequent classification and analysis view the signal trace s(t) from a different perspective. It is desired to reconstruct the basic signature of the original source w(t) given the received seismic trace s(t). Additionally, other information concerning the location, magnitude and type of source must be determined by using the statistical nature of random noise and by identifying unwanted periodic signals. The filtered seismic trace can be approximately reduced to:

$$s(t) = w(t) * r(t)$$

Again, if the earth's reflection coefficient's r(t) could be accurately modeled, de-convolution could be used to recover the original source w(t).

Seismic events are classified into several broad categories. These categories of seismic classifications include [26,27]:

Table 1 Seismic Event Classifications

Natural events:

     tectonic
     volcanic
     collapse earthquakes
     ocean microseisms

Man Made:

     Controlled
     explosions
     cultural noises
     induced
     reservoir impounding
     mining
     quarry
     fluid injection

All of the above classifications are broad and it is often difficult to distinguish between similar events. Much of the ongoing seismic event classification research limits the discrimination goal to that of a bivalent, or two class decision.

The actual classification of seismic events is not a firmly developed procedure. There is no one single method that provides 100% accurate results [14,21,23]. Some of the problems encountered in classification of seismic signals range from inaccuracies in the actual data collection to the shades-of-gray type problem in differentiating between a mining explosion and a quarry blast. The instrumentation used to record seismograms vary in their range of frequency response, which can be on the order of 0.01 Hz for the low cutoff, to 1 KHz at the high frequency cutoff. Typically, data from the Center for Seismic Studies used in this research had a sampling rate of 40 Hz with a low cutoff frequency of approximately 0.5 Hz. If the classification scheme

utilizes only data from a single recording station, the sampling rate and frequency compensation can be applied uniformly without further consideration. Characteristics of the seismic signal will be modified as a function of distance from the event focus to detector location. The distance function alone causes significant difficulty in event classification. Frequency content is attenuated and the signal can potentially undergo modal transformation in which transversal waves will transpose into compressional waves and even possibly reverse again before arriving at a detector.

Automatic means for the detection of earthquakes has been an active goal of seismic research for more than 25 years. Various discriminators have been developed and perform with a high degree of accuracy Detectors based on the ratio of short term signal average to long term signal average, STA/LTA, approximate a Neyman-Pearson filter and tend towards optimal signal/noise ratio [33]. Other detection schemes have been developed ranging from Freiberger (1963) to the SRO (1983).

Recent work within seismology extends the detection problem to include classification of seismic events. This classification involves discriminating between natural seismic events such as tectonic, volcanic and collapse earthquakes verse's man made events such as controlled explosions, cultural noise, and induced events. Knowledge based recognition systems developed by Roberto and Chiarutti use a knowledge base blackboard scheme to automate seismic signal analysis [26].

Neural networks have been studied by Dysart /Pulli, Lacosse, and Dowla/Taylor/Anderson [2,3,10,11]. Most of the neural network research efforts capitalize on seismic signals in which the phases of the signal can be separated and the subsequent parametric data applied as input to a neural network. Results in most cases are very favorable, often with better than 90% classification accuracy [10]. Analysis of parametric data is often done in an off line setting and relies on the judgment of a seismic expert to determine phase information and other parameters before processing with neural networks. Lacosse is concentrating in development of seismic phase identification using neural networks as a supplement to the existing IMS system as part of an ongoing research contract with ARPA's Artificial Neural Network Research Program [2,3].

Seismic Network Analyzer is an expert knowledge based system developed by Vito Roberto and Claudio Chiaruttini [26] that adopts the blackboard problem solving paradigm. The system consists of four basic units; user interface, a permanent database, data and symbolic memory, and knowledge supervision module. A prototype has been demonstrated using seismic data from the Seismological Network of North-Eastern Italy.

1.3   Software Implementation Issues

There are many competing languages that could be considered for implementing a seismic event classification system. The research presented within this paper utilized the ADA language as the primary programming language. ADA is a procedural language specifically developed by the United States Department of Defense (DOD) for use with embedded computer systems [1,28,29]. The language features modern programming concepts such as separate specification

and implementation portions of code, strong data typing, tasking, strict compiler requirements, library management, and other features. Originally, ADA was trademarked by the DOD and the content of the actual language was strictly controlled. This control has since been relaxed but subsets or dialects of ADA are not commonly found as in other languages [28]. Validated versions of ADA maintain a high degree of portability between host machines.

Perhaps the strongest feature of the ADA Language is that it was specifically designed to meet the needs of software engineering. ADA encompasses the entire life cycle of a software project from the initial set of specifications, to design, testing, and ongoing maintenance. Overall, ADA supports and promotes software engineering practices not commonly found in traditional AI languages [6,7,30,32].

Many stand alone AI systems have been developed in the C language, which in itself implies that many of the issues concerned with embedding AI within an ADA environment have already been approached [7]. These existing systems in C, coupled with the software engineering features of ADA, make a strong marriage between AI and the ADA language. ADA is a procedural language that offers a good environment for processing time series data such as seismic wave forms. Neural network functions can be easily written in the ADA language thus combining a model free AI language (neural network) within a procedural language.

The software for the project was written in ADA using the Meridian Corporation's validated compiler. The development and target hardware was the IBM PS/2 model 70. The software developed for the project consisted of two major parts. The first was the development of a complete user friendly package, Seismic Waveform Analysis Package (SWAP), to perform analysis of seismic waveforms using neural network technology. The second part was the implementation of the stand alone neural network programs designed to operate in protected mode.

SWAP consists of a desktop, a top menu bar with pull down submenus and a status bar at the bottom, showing the location of the waveforms and data files. The techniques included in SWAP are the examination of seismic signals, selection of seismic signals, feature extraction of the selected waveforms, the training of a neural network and the classification of a different set of selected waveforms.

Significant characteristics of the waveforms were extracted from the information at the Seismic Center and kept as an index into the raw waveform measurements. The user of SWAP can view these characteristics and select desired waveforms to include in the training/classification set to be presented to the neural network. The user also has the ability to view a plot of the raw data to better determine the appropriateness of its inclusion into a data set.

Once a dataset has been selected, the data can be filtered or preprocessed to perform transformations such as Sine, Fourier and Haar transforms on the signals prior to presenting them to the neural network. This allows the user to perform feature extraction and noise elimination in the set of waveforms. Once the feature extraction is completed, a neural network can be trained using these waveforms. The networks included are Backpropogation Supervised Kohonen,

6

Unsupervised Kohonen, ART-2, and Radial Basis. Upon the selection of a network, the user can enter the initial configuration of the network. The type of information requested will depend on the network selected.

Since network training is such a lengthy process, SWAP was written using ADA's feature of concurrent programming, tasking, which allows multiple processes to share the resources of a single CPU. In this way the training sessions could proceed "in the background" while the end user could perform other duties and investigations on other sets of waveforms. The only limitation imposed was that two training sessions could not proceed at the same time. After training was completed, the user has the ability to save the desktop to a disk file. The information saved includes the waveforms used, the preprocessing algorithms, the choice of networks, the initial network configuration and the weight vectors of the internal layers of the network. This gives the user the flexibility to continue training (with the same or different network topology) or to classify from a set of input waveforms.

Individual waveform raw data and perform filtering or preprocessing can be selected as options in order to better examine the effects of a particular noise reducing technique. As security and recovery methods Seismic Waveform Analysis Package has a set of utility functions which perform automatic Snapshot backups and a log feature which records a user's session. The user can change the time intervals between snapshots and can toggle the logging procedure. During tool development, feature extraction methods found better success with large internal network topologies. This necessitated the conversion of the product to a protected 32-bit mode of operation. The first steps were to port the neural network training and classification to the new requirement of the 32-bit mode.

1.4   Research Plan

This research effort focuses on the viability of using neural networks to classify seismic events using only parametric data automatically extracted from the original seismogram along with the official classification as determined by the Center for Seismic Studies. In contrast to existing knowledge-based systems, this method is not based upon seismological expertise. Parametric wave form representation requires that the essential characteristics of a particular event type are adequately represented by the fit vector presented to the processor.

The seismological aspects of this research could potentially require extensive background training within the field of seismology. By approaching the seismology problem as a signal classification problem, as opposed to that of a purely seismic problem, familiarity with seismic phase identification, travel times and related considerations can be somewhat over looked. The carefully constructed data base used in this research, allows efforts to concentrate mainly on the application of neural networks to the solution of the problem. This data base includes only seismic events that have been analyzed by seismologists and are considered to be correct in terms of parametric data and event classification.

7

Seismological background was investigated to provide a basis for interpreting results, suggest parametric transformations used in other classification schemes, and to provide some heuristics to enhance overall system performance. Actual supervised training of the resulting neural network models only rely on presenting the wave form data along with the correct seismic event classification.

An artificial neural network is incorporated as part of hybrid software simulation system capable of detecting and classifying seismic events. The hybrid system model is composed of

1) classical filtering techniques (signal Pre-processing),
2) neural network (pattern detection and discrimination) and,
3) a rule based system (final pattern classification and pre-processor adjustment).

The optimal role of the neural network is initially assumed to be that of seismic detection and discrimination. Further investigation is proposed during the development of the hybrid system to determine the extent to which the pre-processing (filtering) and post processing (rule based system) can be replaced by the neural network. Additionally, fuzzy logic will be investigated as applied to seismic processing.

1.5 Organization

This introductory chapter, Chapter 1, has offered motivation and a somewhat broad description of the seismic discrimination problem. Some of the current research methods for seismic discrimination was discussed leading to the incorporation of neural networks for seismic event classification. The review of literature indicates that the current state-of-the-art in seismic discrimination is the active utilization of neural networks. The data base used for testing discussed throughout this paper is described in chapter 2.0. The various tables listed in Appendix A with seismic wave form names, stations and Julian dates are sufficient references such that anyone accessing the on-line data base at the Center for Seismic Studies can retrieve the related seismic wave forms. The related software tools developed in later chapters, were implemented in the ADA language. Software implementation issues were discussed in section 1.3.

Seismological background is covered in Chapter 3. The broad classification of seismic events as used by seismologists is presented along with plots of sample wave forms. Qualitative assertions and heuristics that are commonly used for seismic event classification are discussed. Overall strategies and numerical processing schemes in use are summarized.

Chapter 4 discusses seismic parametric conversions. Parametric data is derived from the sampled wave form and is independent of the identification of various seismic phases associated with most classification schemes. Most of the parametric data was derived from sonograms and moment feature extraction Some of the neural networks used for classification is presented in Chapter 5. A summary of the work with a radial basis function neural network is presented along algorithm development found in the Appendix.

2.0   SEISMIC DATABASE

2.1   Overview

The Center for Seismic Studies (CSS) is an agency funded by ARPA with the principle objective of providing the research community easy access to seismic data. Since 1982, CSS has been improving the earlier teleseismic database procedures and programs of the Lawrence Berkeley Laboratory and the Discrimination Group at Lincoln Laboratories. A more progressive database was needed to meet the standards of the seismic research community and an interactive method was needed to access the database. In 1987, the version 2.8 Database was released adhering to the Intelligent Array System (IAS), a type of seismic data collection standard. The Version 2.8 database also embedded SQL to interactively access the seismic database. In 1989, CSS modified the Version 2.8 database to handle regional as well as teleseismic events. The modified database, Version 3.0, also has a simple database structure that was less complicated for the interactive use and lessened maintenance.

The Seismic Operations LAN (SOL) is the primary host for interactive analysis from the seismic research community. SOL is also automated to collect and process external seismic information from various international seismic stations. Using the processing power of a SUN workstation, SOL is the heart of the interactions of CSS to the seismic community. The Central Data Repository (CDR), the seismic data archives of CSS, is the storage facility for SOL. The CDR consists of a 600 Gigabyte Tape drive dedicated to waveform storage, a 6 Gigabyte database management system, and a 400 Gigabyte Optical Jukebox to store satellite imagery, map graphics, and waveform segments. Figure 3 displays the current configuration at CSS.

Currently, CSS is upgrading the Central Database Repository with larger optical drives as well as larger hard-drives with faster SUN workstations to give the research community with more computing power and storage capability.

2.2   Databases at the Center for Seismic Studies

Although the Center has many databases consisting of seismic data that has been collected worldwide, the three major databases are the GSETT, the IMS, and the EXPLOSION. These three databases are 75% of the entire parametric and waveform data stored at the Center.

9
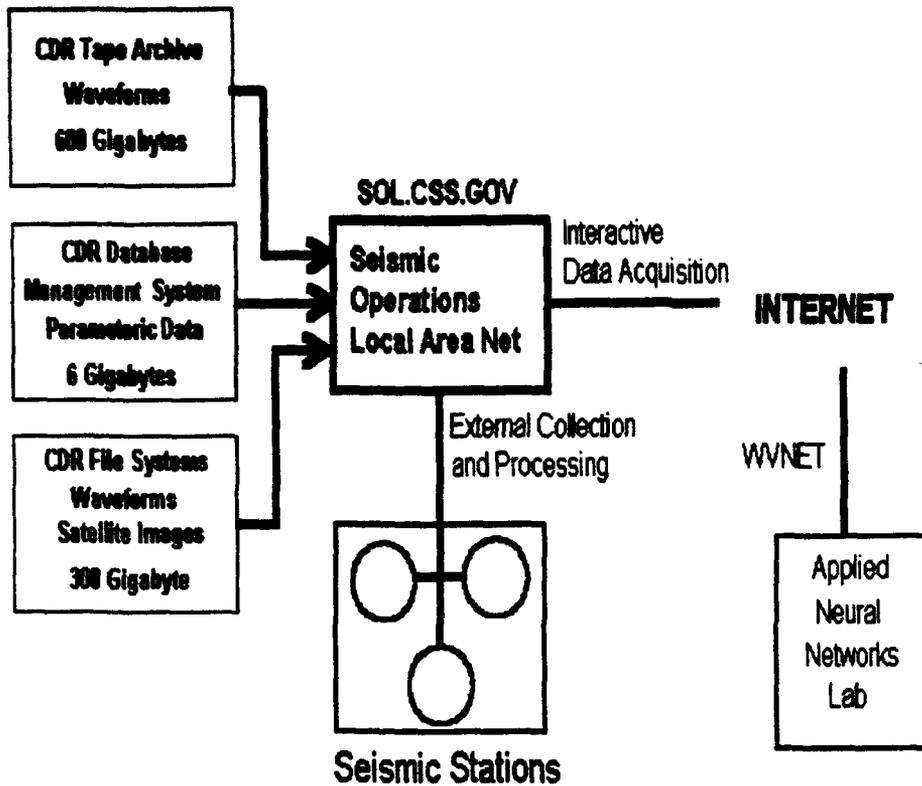
Figure 3 CSS Database

The GSETT database was the work of the *Ad Hoc Group of Scientific Experts to Consider International Co-Operative Measures to Detect and Identify Seismic Events*, called GSE. GSE was formed in 1976 by an international group of scientists during the Conference on Disarmament for the sole purpose of exchanging data useful for monitoring a limited or comprehensive nuclear test-ban treaty. Using approximately 50 international seismic stations, GSE conducted the first international exchange of seismic data in 1986 during the GSETT-1 test. Due to the complexity and size of the exchange of parameteric and waveform data, the test was only a limited success. Waveform data were to be available on request, but never exchanged routinely. But with the increasing technology and the availability of larger computer networks, the second international full-scale test began the 22nd of April 1991 to the 2nd of June 1991. During these 42 days of seismic activity, over 3,700 events were classified and 85,000 waveform

segments were collected and stored into 1.2 Gigabytes of information. Although, the second international test had some small procedural problems, the test was a smashing seismological success.

The Intelligent Monitoring System (IMS) is a ARPA-sponsored computer system for automated processing and interpretation of seismic data recorded by arrays and single stations. It was integrated into CSS computer systems, and has been operational since 1990. The IMS data has been cataloged in the IMS database at CSS, which contains seismic traces from the two largest seismic stations in Norway, ARCESS and NORESS ARRAYS.

The EXPLOSION database consists of all unclassified seismic data on nuclear testing. Another database currently being investigated is the GROUND TRUTH database, created by Lori Grant at CSS. This database is currently being compiled from both the IMS and GSETT databases for the sole purpose of seismic discrimination by neural networks or fuzzy logic. The GROUND TRUTH database consists of a hand picked group of seismic events that were verified through means of seismic bulletins, mining records, and personal contact. Although the database has been released to the public, the numbers of events are not large enough to create a test bed for our current networks.

2.3   Applications at the Center for Seismic Studies

The heart of database management at CSS is the SQL/ORACLE database host. This gives users an interactive method of accessing data. Since SQL querying can be quite taxing, CSS has created some tools making the collection and examination of data easier. To make the seismic tools accessible from many different operational platforms, CSS programmed the tools to be used as Xwindows applications.

CENTERVIEW was the first programmed tool from CSS. Using this tool, one can directly access the database without using the burdensome SQL queries, and still have the power to select the data on a variety of constraints. With this program, one can compile data for downloading, review parameteric data, and transfer data to the other seismic tools. The next tool was *MAP*. This tool displayed the location of the seismic events [epicenters] and the location of the seismic stations that recorded each event. These locations can be displayed on a variety of geographic maps stored at CSS by using the *MAP* program. The last tool created was *GEOTOOL*. This tool gives researchers the ability to view the waveform in a time series plot, seismogram. It also has some signal processing capabilities such as FFT's, filtering, spectrogram, and others.

2.4   Research Databases

Research databases consist of a subset of the GSETT and IMS database, called *SUBSET 1* and *SUBSET 3* respectively. Subset 1 constructed for training purposes for retrieving data from CSS and initial software configuration. The database consists of 75 waveforms recorded in the

11

Euro-Asian Area with a fixed wavelength of 2400 samples and a sample rate of 20 Hertz. This was the preliminary test data set for neural network training. Each event classification was verified through the *REMARKS* database table.

Subset 3 is a waveform set based on the work of Thomas J. Sereno and Gagan B. Patnaik from the paper entitled "Data to Test and Evaluated The Performance of Neural Network Architecture's for Seismic Signal Discrimination." This was a two year study which focused on producing data sets for neural network evaluation. The waveforms selected for use in this particular subset were taken from data set #1 of Dr. Sereno's paper. The data for these waveforms was obtained from the NORESS and ARCESS arrays located in Norway, which consist of 25 short period instruments configured in four Concentric rings with a maximum diameter of 3 km. The data for these waveforms were digitized at a rate of 40 Hz, with a digital gain of 100000 digital counts/volt. The instrument response for these arrays is approximately flat to velocity between 2 and 8 Hz.

This dataset was subdivided into eleven smaller databases. Origin identification numbers (Orids) were selected from among these databases for use in the creation of subset 3. Five Orids from each database were selected and utilized in a query to the Center for Seismic Studies, where the waveforms are available on-line. The initial query provided a multitude of waveforms which provided the basis for subset 3. This pool was further narrowed to 124 waveforms by selecting only those with cb. channel.

The 124 waveforms left after the narrowing process were then downloaded to our location using CenterView. This formed out "Subset 3" database. The waveforms consist of 16.8k data points sampled at 40 Hz, cb channel only, and a 0.006837 calibration.

3.0 SEISMOLOGICAL BACKGROUND

The various aspects of seismology include observational seismology, instrumental seismology, theoretical seismology, and data analysis of seismic events. The primary focus of applying fuzzy logic to seismology was the analysis and subsequent classification of seismic data. Some introductory terminology as applied to analysis of seismic data will be reviewed.

3.1 Overview

The types of seismic events can be roughly divided into two categories: natural and man made [20]. Natural seismic events include tectonic plate movement, volcanic activity, collapse earthquakes, and oceanic microseisms. Man made seismic events can be the result of a controlled event or that of an induced event. Controlled events are typically explosions and cultural noises while induced events will result from reservoir impounding, mining, quarry and fluid injection. Table 2 lists the broad categories of natural and man made seismic events.

Seismogram interpretation is dependent on the location of the recording station and the type of structural model utilized for wave propagation in the geological region of the recording station. The structural models and propagation paths have lead seismologist to three different categories of seismic events, without regard to the source of seismic activity. These categories are based on distance between the source epicenter and the recording station. It is common practice to use a spherical model of the earth and express the distance from seismic event focus to the recording station as the angle subtended at the center of the earth between the focus and the station ($1^o = 111$ km).

The categories thus established are:

Local events $< 10^o$

Regional events 10° to 20°

Teleseimic $> 20^o$

Raw seismograms are relatively lengthy. Typical sampling rates vary between 20 Hz to 40 Hz with high frequency instruments operating at sampling rates upto 1 KHz.. The duration of seismic events range from a few minutes for discrete events to day for seismic swarms. Seismograms used in this research all result from discrete events sampled at 20 Hz, with a total of 2400 data points per sampled waveform. Waveforms were taken from the GSETT database at the Center for Seismic Studies. Figure 4, shown below, illustrates a typical quarry blast while Figure 5 is a typical marine explosion. In each case, the start of the seismic event occurs at sample number 600. This starting alignment represents a 30 second pre-event leader and is common for all seismic traces used in the GSETT database.

Table 2
Types of Seismic Events

Natural events:
                     tectonic
                     volcanic
                       collapse earthquakes
                       ocean microseisms

Man Made:
                     Controlled
                       explosions
                       cultural noises
                     Induced
                       reservoir impounding
                       mining
                       quarry
   fluid injection

13
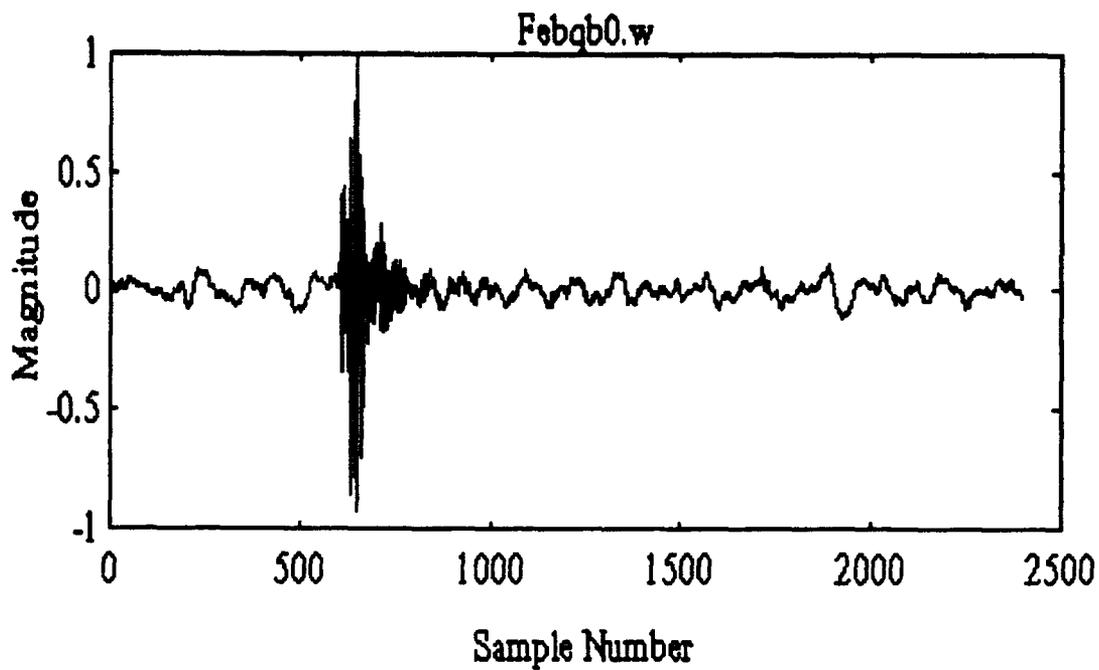
**Figure 4** Quarry Blast Febqb0.w from GSETT Database
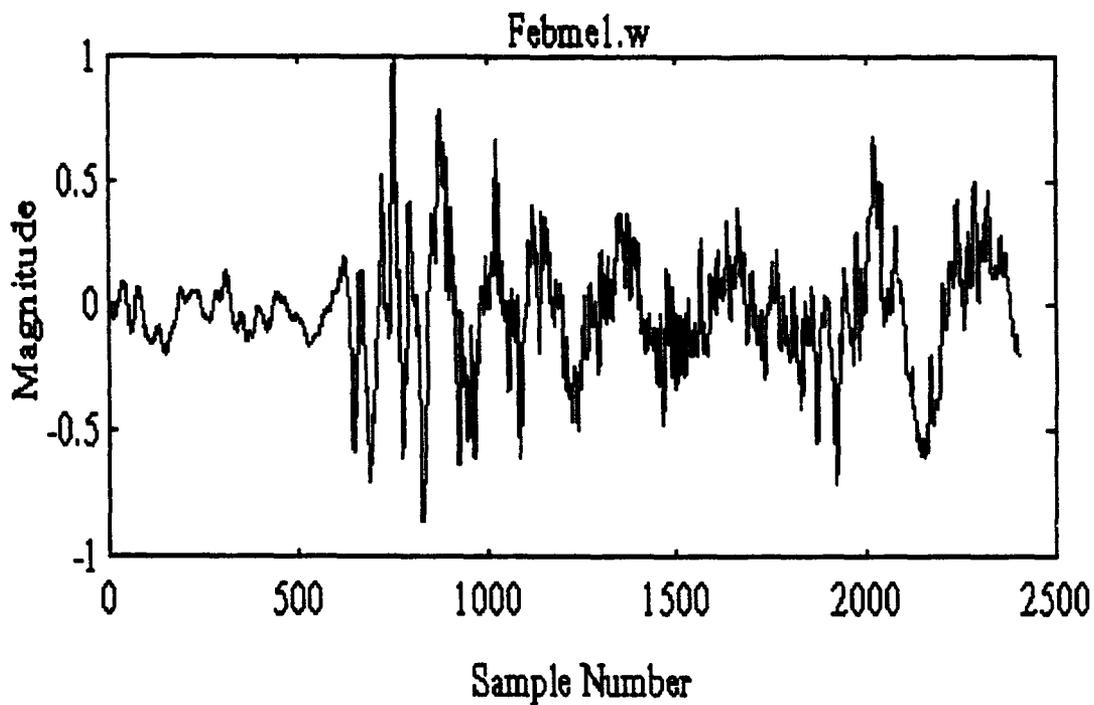
Figure 5 Marine Explosion Febme1.w from GSETT Database

14

In analyzing waveforms such as those presented in Figure 4 and Figure 5, seismologists will identify different phases of the seismogram based on the time of arrival and the mode of propagation through the earth.

There are two basic types of seismic waves [19] , body waves and surface waves. Body waves are radiated by the seismic source and propagate in all directions while surface waves are concentrated along the surface. Body waves can be further subdivided into compressional (longitudinal) and shear (transversal) waves. Compressional waves are often called primary waves or P waves and transversal waves are called secondary or S waves. P waves tend to travel at a rate 1.7 times that of S waves and are normally the first portion o the seismic waves to be present in a seismogram [19].

Figure 6 illustrates the relative motion of a seismic wave and the mode of propagation. The P waves are always the first waves to arrive [19,26]. The P waves are surface waves that cause the rock particles to oscillate back and forth in the direction of propagation and can be compared to the propagation of sound waves. S waves cause rock motion perpendicular to the motion of P waves and represent a shear wave. Motion of S waves through the liquid parts of the earth's interior is not possible since liquids do not sustain shear forces. Two additional waves often associated with a seismic event are the LQ and LR surface waves. The L stands for long , Q represents Love waves and R is Rayeigh waves. These two waves are often dominate in terms of relative amplitude. Love and Rayleigh waves exhibit velocity dispersion which can be observed as frequency variance where-as P and S waves tend to be velocity invariant.

The P, S, LQ, and LR, portion of the seismic trace are referred to as phases. These phases are further subdivided to give indication of propagation path. A Pn or Sn phase indicates a path that is in the upper crust and is confined to the granitic layer. Reflection of phases are possible off other layers in the earth. A phase reflected off the Moho layer is referred to as a PmP or SmP phase. Many other combinations are used as dictated by the seismic event being evaluated.

3.2   Analysis of a Regional Seismic Event

A regional seismic event from the GSETT data base is now presented to illustrate the type of parametric information determined by a seismic analyst. Data base notation as assigned by the Center for Seismic Studies is utilized in the seismic event description that follows. The regional event considered is illustrated in Figure 7. The event is assigned an origin identification within the GSETT data base of ORID = 36907. This event occurred on April 28th, 1991 [Julian date of JDATE = 1991117], and was determined to be a regional event. A summary of the seismogram analysis is given in Table 3.

15
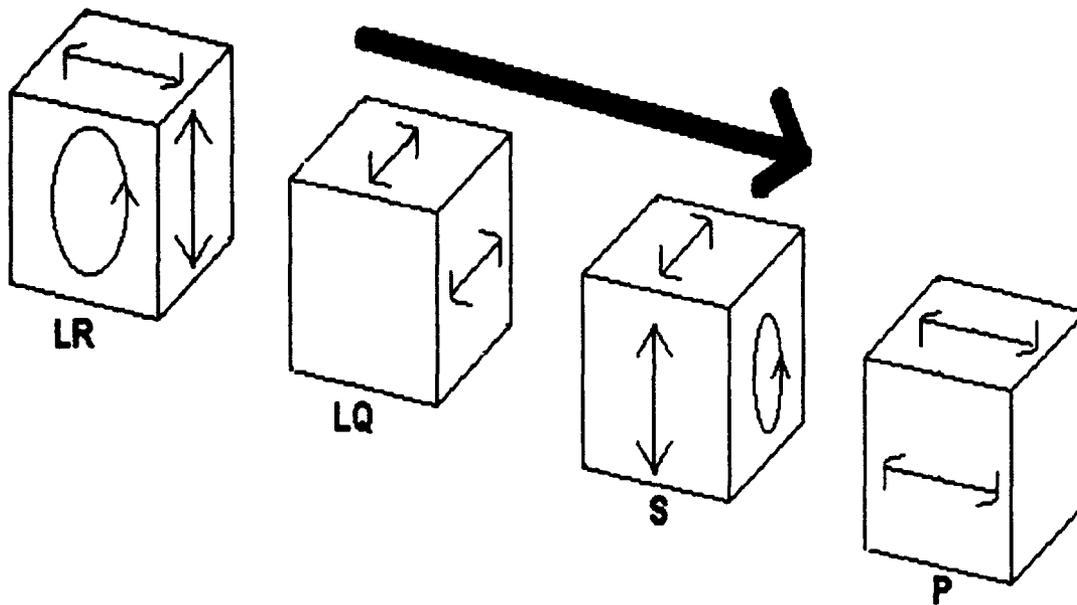
Figure 6 Rock Particle Motion

The STASSID label represents a station association identification number assigned as part of the data base record. The wave train of a single event may be made up of a number of arrivals and the STASSID allows arrivals believed to have come from a common event to be joined together in the data base.

The signal amplitude is denoted AMP and represents a zero to peak amplitude of the earth's displacement in units of nanometers. The duration of a particular phase is designated PER and is in units of seconds.

Figure 7 is a regional event with three recorded phases. The magnitude scale was normalized to +/- 1 with actual displacement magnitudes indicated in Table 3. The first arrival wave is the Pn wave that traveled through the earth's crust from the epicenter to the recording station. A secondary surface wave, Pg, arrived from a deeper propagation path followed by a large magnitude LQ or Long-Love wave. The first 618 sample points (approximately 30 seconds) before the arrival of the Pn wave is a period of no seismic activity. This represents normal background noise and will tend to drift in magnitude throughout the course of the day due to cultural noises.

The recording station for this particular wave form was located in Boyern, Germany. It was recorded with a single vertical channel that measures earth displacement. Table 4 gives the station location and instrument calibration factors. The frequency response of the instrument is plotted in Figure 8. The 3 dB bandwidth is 3 Hz. A usable bandwidth of about 10 Hz can be created with appropriate inverse filtering of the seismic waveform.

16

TABLE 3

Seismic Analysis of Regional Event FEBR9.W

ORID	36907
Date	April 28, 1991
Julian Date	1991117
Event Time	672777893.300 seconds from January 1, 1970.
Classification	Regional event
Recording Station	Grafenberg Array, Boyern, Germany (GRA1)

Event Location

Latitude	46.22°
Longitude	15.44°
Depth	8 Kilometers

Phase Information

    3 phases recorded at GRA1

    Surface Wave Magnitude measured at 2 nanometers

    Body wave Magnitude measured at 3.50 nanometers

Phase Summary

Phase	Start Time	Start Sample number	ARID	STASSID	AMP	PER
Pn	672777957.3	619	492530	368441	41.2	0.65
Pg	672777971.3	886	492531	368442	323.6	.082
Lg	672778033.8	2136	492532	368443	468.0	0.71

17
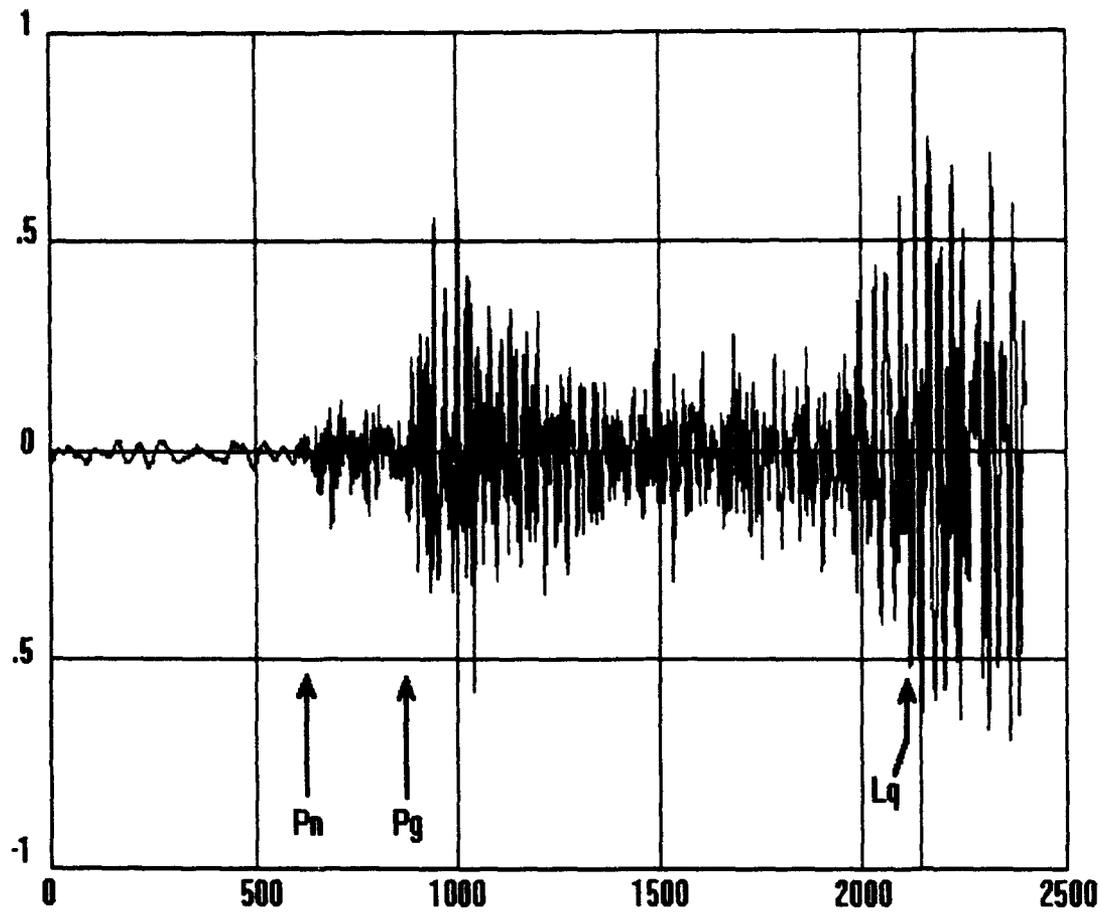
Figure 7 Regional Seismogram FEBR9.W

Table 4
Station Information

GRA1 - Grafenberg Array -- Boyern, Germany

> Single Station
> Channel Type: bz
> Channel Id: 51671

Location

> Latitude   49.692°
> Longitude   11.222°
> Depth   0.5 Kilometers From Mean Sea Level

Noise Measurements - Correction Factor

> Mean Noise -   6.5 nM
> Stand Dev   -0.2 nM
> Signal to Noise Threshold   1.5
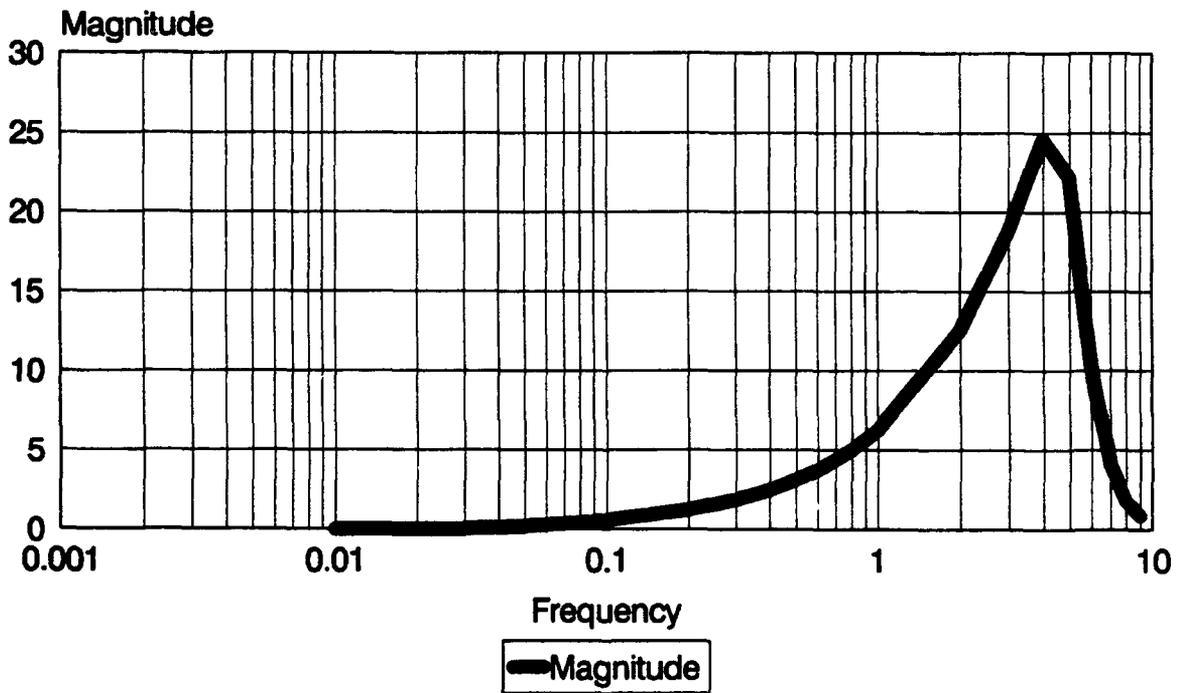
Figure 8 Frequency Response of Grafenberg Array Channel bz

3.3 Qualitative Assertations and Heuristics

The seismic analyst evaluating a given wave form must base his reasoning on a physical model of the earth with respect to the recording station location and the suspected seismic epicenter. Qualitative assertions must be made concerning the propagation in a global scale. Several qualitative assertions made pertaining to seismic events are listed in Table 5 [17,18,26,27]. These assertations are based largely on the identification of seismic phases.

Table 5

Qualitative Assertations

1. The dominant frequency of the seismic signal is inversely proportional; to the distance of the event.

2. The Pg wave is the first arriving wave for local events,
   Pn for regional events
   P or PKP for telesiesmic events.

3. The longer the duration, the greater the magnitude.

4. Presence of a strong S-wave is a distinctive feature of   natural events such as earthquakes.

5. The absence of S-waves or weakness with respect to P   waves indicate an explosive or artificial seismic source.

6. Similar waveforms are present in seismograms that originate in the same seismological area.

The assertions listed above are supplemented by heuristics developed by seismologists. Many of the heuristics can be utilized as linguistic descriptors in the development of a neural network seismic event discriminator. Table 6 lists some of the heuristics [15,24,26,27].

Table 6

Seismic Heuristics

1. If the duration of a signal is less than one second, it is most likely noise.
2. If two different signals have dominant signals whose ratio is
   above 10 , then they probably belong to two different events.
3. If the dominant frequency of the first arrival is above 7 Hz, then the
   seismogram belongs to a local event.
4. If the dominant frequency of the first arrival is between 2-7 Hz,
   then it belongs to a regional event.
5. If the dominant frequency of the first arrival is below 2 Hz then it
   belongs to a teleseismic event.
6. The beginning of a seismic event can be detected using Dixon's test
7. Cultural noise will have dominant frequencies above 1 Hz.
8. Microseismic events will exhibit low frequency broad band noise
   from less than 0.01 to 0.5 Hz with periods of 2 to 100 second.
9. P wave is normally recorded first.
10. P is normally followed by S,LQ,and LR.
11. P waves have linear polarization.
12. LR will have elliptical polarization.
13. Earthquakes produce approximately equal amounts of P and S waves.
14. Explosions produce more P waves than natural events.
15. Earthquakes give anaseismic and kataseismic first onsets.
16. Explosions give anaseismic first onsets everywhere.
17. Earthquakes have relatively deep foci.
18. Explosions have shallow foci.
19. Duration's of wave trains are shorter for explosions than for earthquakes

Most of the qualitative assertations and heuristics are based on the various phases of a waveform as identified by a seismologist. The listed assertains and heuristics offer several clues that aid in the development of a neural network parametric conversions. In particular, the heuristics dealing with dominate frequency are examined in Chapter 4.

3.4   Discrimination Methods

Many techniques of discrimination [33] have been used over the years. Various techniques include; amplitude ratios [8], spectral properties, ARMA process model, sonogram detector [17], time independent structures.[17], knowledge based systems [25]. spectral modulation [13], neural networks using spectral data, and neural networks using cepstrum variance and amplitude ratios [11].

21

In all cases, a generalized strategy is used by the seismic analyst. Trace segmentation is used to isolate independent events from the seismic trace and different types of feature extraction methods are employed. A frequently used method is to filter the wave form into three frequency bands. These bands loosely fall into the 0-2 Hz, 2-7 Hz and 7-10 Hz ranges. Division of the seismic trace into these bands often simplifies phase identification. The extracted features are then examined for clarity. This helps establish whether the final analysis is clear, probable, or possible. At this point, a working hypothesis can be formed which will lead to a refined set of calculations and cross checking with other recording stations solution.

One of the most extensively used technique of event discrimination is based on the amplitude ratios of different wave groups [31]. An example of theses method is given in a case study conducted by Wuster for discrimination of chemical explosions from earthquakes. Seismic data was divided into 4 time windows of 10 seconds duration each. The first window contained noise preceding the onset of the event. The second window typically contained the P phase, third window containing a S phase and the forth window with a surface wave, possibly R phase. Ratio amplitudes of each window are formed and discriminate plots constructed with the training data set. Discriminate functions are then determined. Mis-classification percentages of the case study data set where typically less than 10%. Results have not been generalized to a less homogenous data set and are restricted to the bivalent case of chemical explosions verses earthquakes [33].

Research has been under taken in the application of neural networks for classification of seismic events. The modified IMS system described in Chapter 1 incorporates neural networks to supplement the classification process and utilizes phase identification as the main parametric training data [2,3]. Work by Dowla, Taylor and Anderson [10] uses a backpropogation neural network. Phase ratios of seismic events serve as the input to the neural networks. Preliminary results of a bivalent case discriminating underground nuclear explosions from earthquakes have achieved a correct classification rate of 93% [10]. Dysart and Pulli [11] report that wide band spectral ratios Pn/Sn and Pn/Lg provide good discrimination between earthquakes and mining explosions. Using a data set of 95 seismic traces, Dysart and Pulli trained a backpropogation neural network with spectral ratios and achieved 100% classification of the test data set [11].

4.0 SEISMIC PARAMETRIC CONVERSION

The information contained in a seismic trace is somewhat hidden when only the time series wave form is considered. By using various parametric transformations, these wave forms can be made to yield some of the hidden knowledge such as the type of event that originated the seismic trace. The dominate frequency of the first thirty seconds of the trace has been found to be an indication of the relative distance to the events origin. The duration can be a clear signal in distinguishing a naturally occurring event from that of a man made event. Many of these transforms that have been found useful by seismologist, were discussed in Chapter 3. The transformation of raw seismic data into parametric data useful for neural network training and classification is examined in this chapter.

The seismic database wave forms were tested with various transformations and the results presented graphically for visual interpretation. The most useful representations included simple time series plots, sonograms, moment feature maps, fractal dimension and relative power plots. Other transformations used for seismic classification are well documented in literature including scalograms, power cepstrum, and cosine transformations. Many of these transformations yield interesting results but will require a more comprehensive study in future research for application as parametric transformations useful in neural networks schemes.

4.1   Fractal Dimension in Seismic Signal Classification

Examination of various seismic traces suggests a self similarity between successive windowed samples of each trace. Fractal dimension quantifies the self similarity of the graphically presented wave form.

While viewed graphically, the wave form occupies a percentage of the two dimensional graphic space, but does not entirely fill the entire graphic region. A completely filled space (all of the graph colored black on white), would have a dimension of 2. A single line spanning the domain of the graph would have a dimension of 1. The seismic wave form is neither composed of a single line nor does it fill the entire graphic space. The wave form will have a dimension some where between 1 and 2.

Five variations of fractal dimensions were used in determining the usefulness of fractals in the classification of seismological events. The fractal variations are derived from two basic fractal computations; compass dimension and grid dimension.

The compass dimension evaluates the relationship between the magnitude length and the ruler length of the signal as shown in Figure 9. If a segment of a signal has N as the magnitude length and r as the ruler length, the fractal dimension, D, is calculated by the equation

$$D = \text{Log } N / \text{Log } 1/r.$$

The grid dimension superimposes a grid pattern over the signal and evaluates the relationship between the number of grid elements through which the signal passes to the linear number of squares as shown in Figure 10.

The fractal dimension of a seismic trace could be potentially calculated graphically by plotting the wave form and counting the number of pixels it occupied in a $q \times q$ grid. If N is the number of occupied pixels and q^2 the total number of pixels in the grid, then the fractal dimension is given by:

$$D = \text{Log } N / \text{Log } q^2$$

The graphical method of fractal dimension does not lend itself to processing large amount of data quickly. The graphic process requires plotting of the seismic trace with a second pass

through the entire graphic area to sum the number of used pixels. A more direct approach for determining fractal dimension can be derived.

Consider a seismic wave form of 2400 data points normalized to ±1200 instead of ±1. Conceptionally, this represents a two dimensional grid of 2400 by 2400 points and from a graphical standpoint, can be used for direct calculation of the fractal dimension. The scaling ratio r becomes:

$$r = 1/N^{1/2} = 1 / (2400 \times 2400)^{1/2} = 1/2400.$$

The integer distance from one data point to the next data point is summed for a total of the N points (or parts). This is roughly the total length of the wave form.

$$\text{Total length} = 1/N \sum_{k=2}^{N} (1 + (2/N(x_k - x_{k-1})^2)^{1/2}$$

The total length would not exactly represent the number of parts or occupied pixels if a large amount of clutter is present in the wave form. A comparison of the strictly graphical method to the modified method using total length yields no significant difference in fractal dimension when using seismic wave forms. The fractal dimension of the modified grid can be estimated by:

$$D = \text{Log (Total trace length)} / \text{Log (Number of grid points)},$$

which is equivalent to the fractal value as determined by the compass dimension method.

Four variations of the grid dimension method were used for classification. The first variation uses a square window, the number of horizontal and vertical grid elements are equal. The second variation implements a rectangular window where the number of vertical elements is greater than the number of horizontal elements. The third and fourth variations high pass filter the signal before variations one and two are applied.

For each method used, the seismic signal is divided into several time slices, windows, and a fractal dimension calculated for each window. This produced a series of fractal values upon which a neural network was trained and tested for classification.

The neural network has a five neuron output. Each neuron denotes a specific type of event. Since the output neuron values may vary between 0 and 1, the neural network output is processed through a fuzzy rule set to determine final classification. The final results of classification percentages may be seen in Table 7.

Table 7
Fractal Dimension Classification Results

Fractal Dimension Method	Classification
Compass dimension	45.30 %
Grid Dimension - No filter	
Square window	4.00 %
Rectangular window	8.00 %
Grid Dimension - High Pass Filter	
Square window	14.67 %
Rectangular window	16.00 %
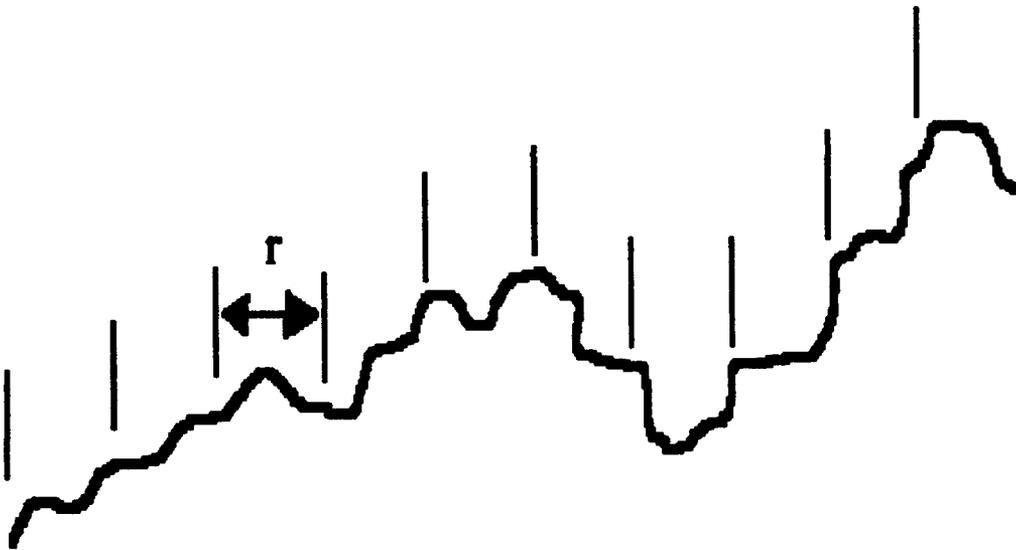
Figure 9 Compass Dimension Method
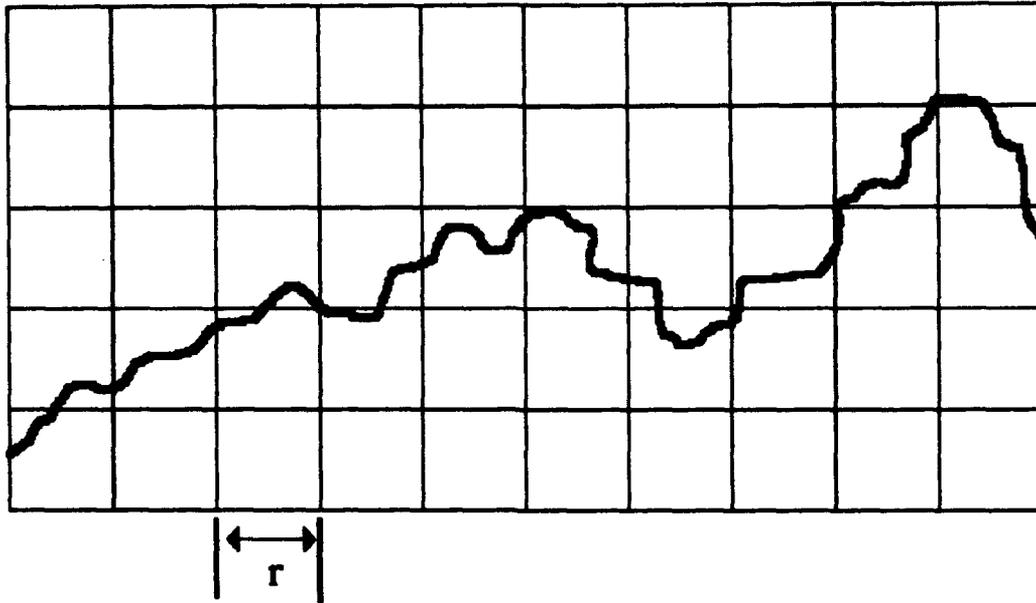
Figure 10 Grid Dimension Method

4.2   Sonogram Feature Extraction

The Sonogram is one method of extracting frequency data for presentation to the neural networks for classification of seismic events.   The first part of the procedure is to normalize the seismic trace by dividing the entire segment by the largest magnitude in the segment.   Then, the seismic trace is "windowed" , divided into equally spaced segments of the original trace size.   For example in Subset1, where all waveforms were 2400 samples long, the trace was divided into 32 different segments.   This produced 32 segments with 75 samples in each segment.   The Fourier transform was taken of each window to created a 3-dimension matrix where the dimension where window, frequency, and magnitude.   This array for Wave 1 of Subset 1 can be observed in Figure 11.

26

Frequency   time
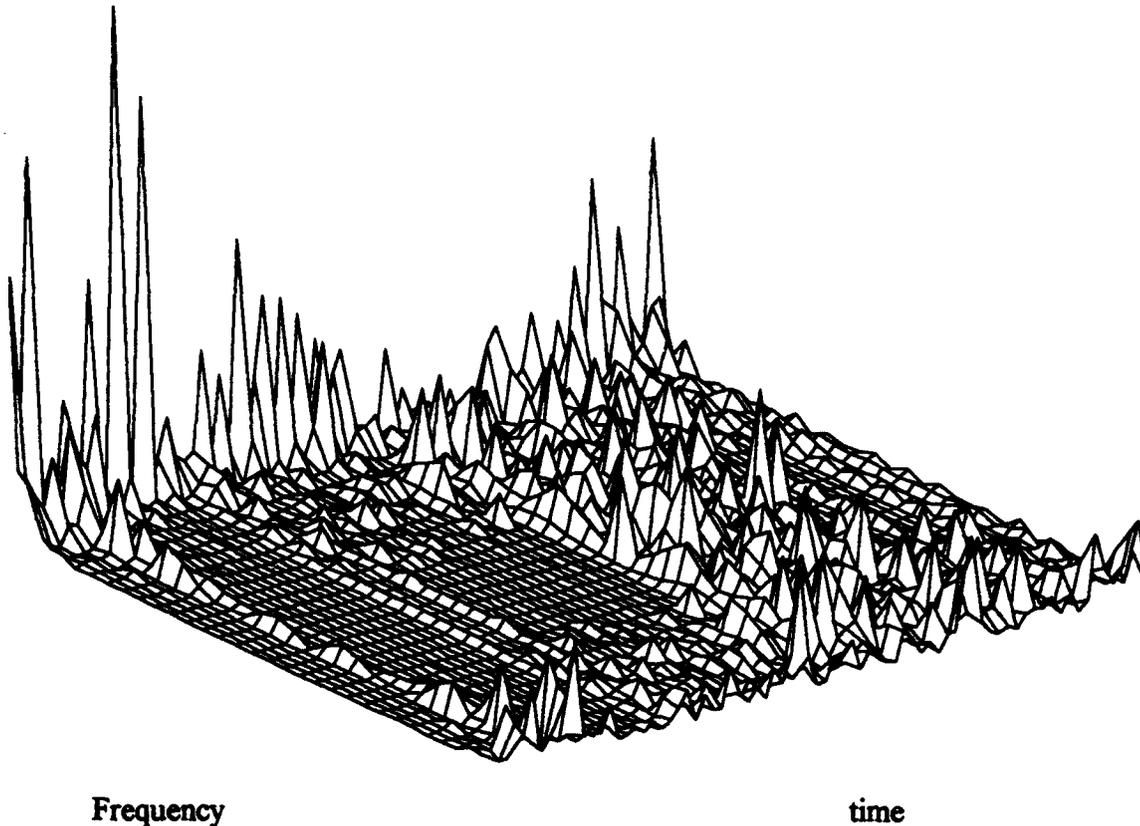
Figure 11 Sonogram of Wavell

The columns of this matrix are transposed and concatenated to form a single vector from the larger matrix. This is performed on a the segment, waveforms, from the database and presented to the network. This created a slight problem since the size of the parametric data was large and cause longer computational time when presented to the network. The routine was extracting too much data.

One method of solving this problem was to have less windows, and another was to choose a method of finding the particular frequency that extract the most information. Using the first method, we found that a 16 windows reduced the size of the data adequately. The second method can be found in the dominant frequency section in this paper.

A backpropagation algorithm was trained on the Subset 1 database with various window sizes. One problem to be noted was that the offset, DC bias, of the waveforms caused some error in the training of the sonogram data. This was due to the magnitude difference of the Fourier transform and the DC offset. Therefore, the waveform mean was subtracted from each segment to remove the offset. This enhance the classification of the data to approximately 87% for the Subset 1 database.

27

4.3   Dominant Frequency in Seismic Signal Classification

Heuristics on seismic signals have presented rules which suggest that the dominate frequency of the first arrival phase is an indication of the event type. The specific heuristics are:

1. Cultural noise will have a dominant frequency above 1 Hz.

2. If the dominant frequency of the first arrival is below 2 Hz, then it belongs to a teleseismic event.

3. If the dominant frequency of the first arrival is between 2 - 7 Hz, then it belongs to a regional event.

4. If the dominant frequency is above 7 Hz, then it belongs to a local event.

The training data set from the Center for Seismic Studies has the start of the seismic event aligned 30 seconds (600 samples at 20 Hz sampling rate) from the start of the seismogram. The first arrival phase is generally considered to be within the first 30 seconds of the event wave train and contains the dominate frequency referred in the heuristics listed.

There is no general agreement in the literature surveyed as to the exacting definition of dominate frequency. The heuristics suggest division of the seismic trace into frequency bands of 0 - 2 Hz, 2 - 7 Hz and 7 - 10 Hz. The data base uses a sampling rate of 20 Hz for a span of 120 seconds. The event is aligned by the Center for Seismic Studies data base manager such that the event start time occurs after 30 seconds of pre-event noise. The dominate frequency as described by the heuristics, is only useful during the first 30 seconds after the onset of the first seismic waves. Only sample numbers 600 through 1200 are in the first arrival window that gives the dominate frequency.

The algorithm used to extract the dominate frequency is given by:

1. Filter the seismic trace into 3 banks of signals with pass bands of 0-2 Hz, 2-7 Hz and 7 Hz to 10 Hz.

2. Calculate the net energy in each band and threshold against some minimum value above noise level.

3. Apply a simple comparison rules to generate grade of membership values for the set:
   {noise, low band, mid band, high band, no clear dominate frequency}

Literature suggests that after the first 30 seconds of any given event, the dominate frequency provides no clear indication of the event type. Only the first 30 seconds after the onset of a seismic event contains useful dominant frequency information.

Currently, the two methods of identifying the dominant frequency of a signal are:

1. Band pass filter the signal and evaluate the power in each band, and

2. FFT the signal and sum the energy in each band.

The resulting mesh plots for these methods are shown in Figure 12 and Figure 13 respectively.

The neural network has a five neuron output to present the class type, one neuron for each class. Each neuron ranges between 0 and 1 so, indeterminate levels may be generated. The training results are shown in Table 8.

Table 8
Dominant Frequency Classification Results

Method	Classification
Band Pass Filter	80.0 %
FFT	88.0 %

The Power vs. Frequency Plot of FEBME1.w
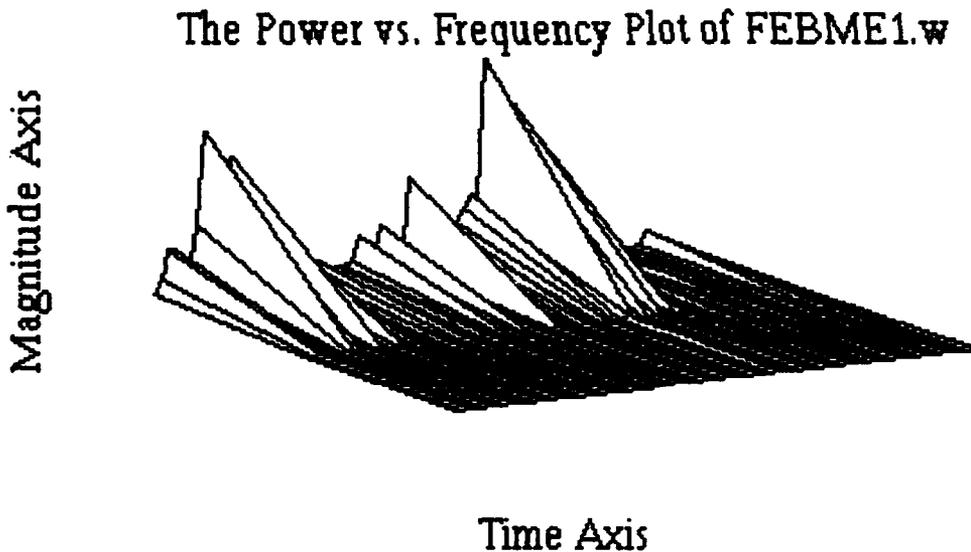
Figure 12 Dominate Frequency Band Pass Filter Fit Vector

The Dominant Feature Plot of FEBME1.W

Magnitude Axis

Time Axis
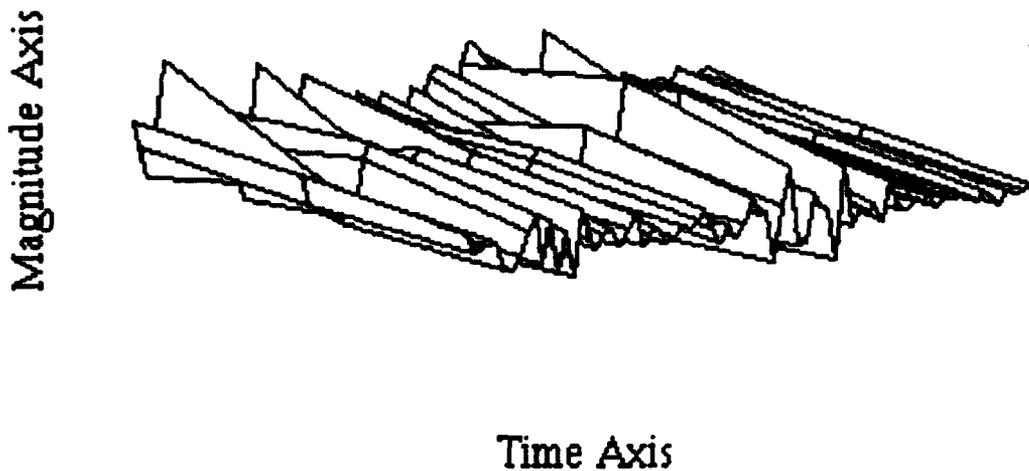
Figure 13 Dominant Frequency FFT Fit Vector

4.4   Moment Feature Maps

One of the rewarding aspects of research is following a wisp of an idea that leads to fruitful results. The calculation of mean and variances are typical signal processing methods used in conjunction with seismology. Bispectrum analysis has been tentatively explored by some researchers and the results suggest that the calculation of higher order spectrums and cumulants may yield interesting and potentially useful results in seismic classification. By following the suggested research, it was necessary to calculate higher order central moments as a prelude to cumulant calculations. Mesh plots of these intermediate results (central moments) produced visually different plots of different classes of seismic events. A key rule of thumb employed, but undocumented by neural network researchers is; if you can visually distinguish different patterns graphically, it's is possible to train a neural network to distinguish the same patterns. Through proper normalization, a moment feature map is constructed with a normalized height < 1 for each window.

The general equation for the calculation of moment features is that of central moments [62].

$$M_n = 1/m \sum_k (x_k - \eta_x)^n$$

where η_x represents the mean value of x.

and   n = moment number , k = sample number.

30

Figure 14 illustrates a moment feature plot of the quarry blast FEBQB0.w. The occurrence of a strong high order moment corresponds to the peak energy of the quarry blast. The right hand side of the plot shows the signal settling down to display wide sense stationarity and possibly strict sense stationarity of the seismic activity after passage of the quarry blast. The production code for moment feature generation is detailed in Chapter 5.
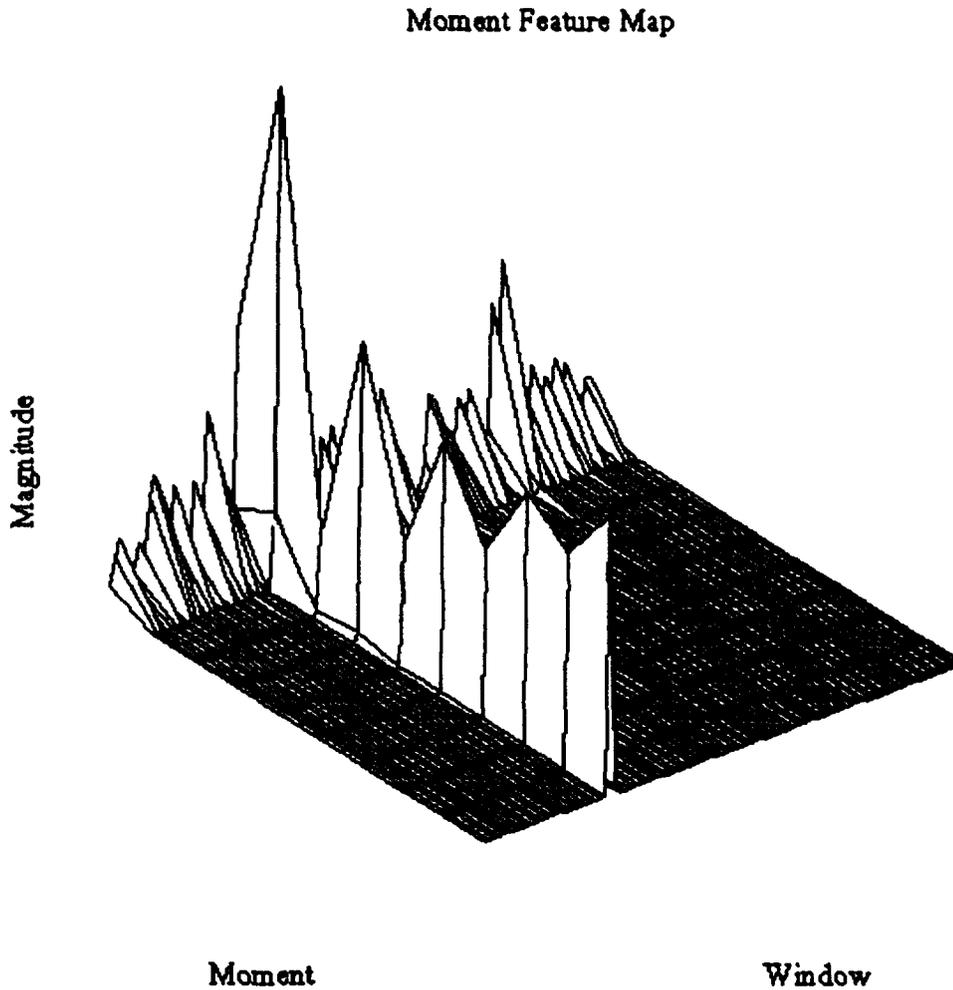
Moment Feature Map

Figure 14 Moment Feature Map of FEBQB0.w

5.0 PRELIMINARY TESTING AND RESULTS

As encouraging as neural networks appear, the high classification rates typically reported are usually limited to the bivalent case and have not been generalized to a multiple class discriminator.

As each waveform is considered an input pattern to the neural network, raw input vector length is prohibitive; therefore, parametric representation is necessary to more succinctly represent the original waveform. This reduced representation is then presented to a much smaller and, consequently, faster neural network. Faster performance yields realistic training times and more reasonable computational system requirements. Ultimately, this is envisioned as an economical PC-based system for preliminary seismic waveform processing and classification.

5.1   Network Testing - Moment Feature

The maximum sample magnitude is used to normalize the other sample values within each individual waveform. Then, time series waveform, consisting of 2400 samples, is divided into a given number of "time slices." Initial research indicates that approximately 16 time slices works well for seismograms of the given length. Evidently, too few time slices prevent adequate resolution of waveform transitions and phase behavior. Specific central moments are calculated for each time slice, the results contained in this paper are based upon calculation of the first ten central moments. Waveform parameterization may ultimately include combinations of FFT-derived spectral components, or fractal-based dimensions; however, only central moment parameterization is considered here. A general form for the central moment calculations is presented immediately below. This relatively simple calculation is represented by the ADA-based algorithm as,

```
--PROCEDURE Find_Moment IS--------------
sum : FLOAT;
k   : INTEGER;
BEGIN
   text_io.put_line("Calculating moments");
   no_slices:=integer(float(length)/float(slice_size));
   for i in (1..no_slices) loop
     sum:=0.0;
       for j in (1..slice_size) loop
         sum:=sum+float(signal((i-1)*slice_size+j));
       end loop;
       mean(i):=sum/float(slice_size);
   end loop;
-- The following loops will calcualte the higher order-- central moments
   FOR i IN (1..no_slices) LOOP
     FOR j IN (1..10) LOOP
       m(i)(j):=0.0;
     END LOOP;
   END LOOP;
   for i in (1..no_slices) loop
     for j in (1..slice_size) loop
       k:=(i-1)*slice_size+j;
       for l in (1..10) loop
```

```
m(i)(l):=m(i)(l)+((float(signal(k))-mean(i))**l)/
                              float(slice_size );
      end loop;
    end loop;
  end loop;
  text_io.put_line("Finished Moments");
END Find_Moment;
_************** **************
```

The above algorithm forms a matrix of row and column dimensions equal to number of moments and number of time slices, respectively. Moment order increases with increasing row number and time increases with increasing column number. A visualization of the processed seismogram appears in Figure 1 (normalized data).

Every seismogram is processed individually, with each waveform's moment data filling in this temporary "moment matrix." The columns of this moment matrix are transposed and concatenated to form a single row in a larger, permanent matrix. Ultimately, the moment data from each processed seismogram appears as a row in this larger matrix—it is important to note that the sequential nature of both the central moments and the time slices is preserved within a row. The permanent matrix is written to disk, in ASCII format, for later presentation to the neural network. Initial testing is limited to a database of 75 seismograms, consisting of five seismic event types, with each type equally represented by 15 examples. Final preprocessing appends the event classifier to the end of each row in the moment data file.network architecture

A back-propagation network is used for event classification based upon the parametric data. A network architecture consisting of an input layer, two hidden layers, along with a one-of-n encoded output layer yielded the most favorable results. The input layer size is matched to the number of datum in a single input pattern; namely (10 moments)*(16 time slices) = 160 elements per moment vector. Neurons per layer, for the two hidden layers are 30 and 24, respectively. One-of-n encoding on the output layer results in five output neurons, a single output neuron for each of the five seismic event classes. Log-sigmoidal transfer functions, with a range of ± 1, were used for all neurons in the network.

Because of the small number of readily available waveforms matching the limited focus of this initial research, training data was limited to using 45 of the 75 available moment vectors. Moment vectors actually used in the network training phase were randomly selected from the parent group. Training time on such a limited number of examples was minimal (typically less than 15 minutes on a 50 MHz 486 PC). However, the convergence rate of the network was impressive even for a small training set and suggests favorable training times on much larger databases. At the conclusion of training, the activation levels for non-target neurons was within four percent of the transfer function zero for all input patterns. The target neuron's activation level was within three percent of the transfer function maximum for all input patterns.
network testing

The remaining 30 moment vectors were used to test the trained network. Because the remaining moment vectors included examples from all five classes of seismic event, the testing phase required the network to accurately identify several examples from the available classes. Output identifiers correspond one of the five classes: local area, regional, marine explosion, quarry blast, and teleseismic. The trained network was able to correctly classify 68% of the waveforms in the testing database—actual activation levels for all output neurons were within four percent of the target value. It must be explicitly stated that rigorous testing of the neural network on large databases has not been conducted; therefore, no estimates of the network's ability to successfully generalize can be assessed at this point.

5.2   Radial Basis Function in Seismic Signal Classification

The radial basis function network performs generalization and discrimination of input patterns using an external teacher for an application of seismic waveform classification. Important modifications to this scheme include (1) change of the size of the spheres; (2) a random walk scheme during testing; (3) the initial radii gradually decreasing to avoid overlap of two distinct regions; (4) a conflict resolution mechanism; and (5) a simple means of decreasing the sphere radius. The applications to seismic signals include using the moments over a sliding window and the first several points of a wavelet. The speed of training of this network exceeds that of backpropagation with the same error rate.

5.2.1.   Radial Basis Overview

In many problems involving pattern recognition, there is an underlying feature space where each dimension corresponds to some measurement of a feature. The number of dimensions is the number of features to be measured. The item to be classified is applied to the network, measurements are taken and the item is mapped to a point in feature space. As items are mapped to points, the points indicate regions corresponding to the classes that are to be differentiated. There are two major issues: (1) Not all possible items will be presented to the system and not all points in feature space will be tagged according to a specific category. The feature space will have holes or gaps that should be filled in with some category indication. This process of filling in the gaps is called "generalization". (2) The boundary between two classes may be very complex. An important assumption is whether the measurements alone are sufficient to unambiguously distinguish items in the classes. If so, then a mechanism must exist to approximate this boundary to any arbitrary degree. Of course, the more complex the boundary (and the class regions need not be connected), then more sample items are needed for distinguishing regions. This process is called "discrimination".

5.2.2   Radial Basis Function Network

The Radial Basis Function (RBF) network proposed here achieves these two goals of generalization and discrimination. It is based on selecting small "sphere" contained in feature space centered on a sample input item[22]. The sphere cam grow or shrink to accommodate

34

generalization and discrimination. A growing sphere is generalizing such that nearby points in feature space are now included in the class associated with the sphere. A shrinking sphere allows for greater discrimination by allowing more volume of feature space to be used by another sphere for another class and by fitting more neatly into a complex boundary. New spheres can be incorporated to the network as more input samples are presented to the network.

5.2.3 Classification

The RBF network consists of three layers: input, middle, and output. For classification tasks the purpose of the input layer is to feed into the Middle Layer (ML) such that each node in the ML receives all inputs. The ML nodes are responsible for the formation of the spheres. A ML node becomes active if the input corresponds to a point in feature space occupied by the sphere of the node. Each ML node has one output that is connected to an Output Layer (OL) node. The OL nodes correspond to the classes or categories. If an OL node receives any active inputs from the ML, then its output is active, signaling the classification of the input.

5.2.4 RBF Learning

The learning procedure in the RBF network is considerably different from other conventional neural networks such as backpropagation. Three main tasks in learning are (1) the incorporation of new spheres by middle layer nodes not previously used for classification, (2) sphere growth, and (3) sphere atrophy. At the beginning of the training phase, the ML has nodes that are not associated with any region of feature space. When an input occurs and no ML node is active, a ML node is incorporated so that it learns the current input and uses it as the center for the sphere. A default radius is assigned to the node so that any feature input that has a distance from the center less than the radius will cause the node to become active. The node must have enough processing ability to compute this distance and perform the comparison with the radius as well as enough memory for the sphere center components and the radius.

Once a ML node has been incorporated into the RBF network by assigning a center and radius, it must make an upward connection to the correct node in the OL on the basis of which category the ML node is associated. An external teacher is needed for establishing this connection.

The learning process involves the modification of the incorporated spheres as more inputs are applied. If an input occurs that corresponds to a point within an existing sphere of the same class as the input, then the sphere radius is increased. If an input occurs within a sphere of a different class, then the radius of the sphere is immediately shrunk to be the distance between its center and the offending point. Once a radius has been shrunk in this manner, it is not allowed to increase at any later time, although it may be shrunk again.

5.2.5   RBF Problems and Fixes

During the training phase it is possible that each ML node simply memorizes the input and the resultant spheres fail to generalize in a useful way. This situation is likely when there are too few sample input items ad they are well separated from each other. In addition, a decision boundary may be very complex, requiring small radii for the spheres and therefore a large number of spheres. Since any software implementation or hardware realization can allocate a maximum number of spheres, they could be all incorporated before the training process is complete. The consequences are that the boundaries are insufficiently approximated and holes exist within feature space that are not associated with any class. Overall, the RBF network looses its performance edge.

A possible solution for this problem is to remove spheres that are completely covered by the union of other spheres of the same category. These removed spheres are ML nodes that can be incorporated later. However, from experimental observations, very few spheres are completely covered this way and should not be removed. If the centers of spheres are allowed to shift positions, then a more uniform coverage of feature space is permitted and then it may be easier to cover redundant spheres and thence remove them. Moving sphere centers is non-trivial since their radii may have to change if the spheres are too close to a boundary.

Another problem results from adding a new sphere with a default radius. It may turn out that the radius is too large such that the new sphere exceeds the decision boundary or covers another sphere of a different class. This problem is mollified by causing the default radius to decrease with the number of input samples. Presumably, as the training proceeds, the boundaries become more well defined and a newly incorporated sphere with a smaller default radius should not perturb the boundary by much.

From the above discussion it is easily seen that spheres of different classes can partially overlap. Then during the classification procedure, two or more OL nodes could be active giving an ambiguous answer. A remedy for this problem is to let the OL nodes have levels of activation based on the number of ML nodes that are active. Therefore, if an input item maps into a point in feature space covered by several spheres, then the number of spheres for each class is counted. The class with the largest number of spheres covering the input point is considered to be the correct class. In a neuronal setting, the OL could have a Winner-Take-All (WTA) circuit to select the class.

Experimental observations indicate that in the event of overlap, usually there are only two spheres of different classes and the above remedy is insufficient. In this case, the remedy is enhanced by considering the center-to-input distance and selecting the sphere and its class with the shortest distance.

The problem complementary to overlap occurs when an input sample is not covered by any sphere. This gap in feature space can occur as a result of learning when a sphere that has been effective in classifying inputs is forced to reduce its radius and can no longer cover the volume it once did. Two solutions for class selection have been proposed. One solution performs

a random walk on the input sample until it falls within the domain of a sphere. Then the class for that sphere is considered to be the class for the input. Another solution is to measure the distances from the input to the boundaries of all other spheres and then select the closest sphere. The first solution requires less computation while the second may give the better answer.

Interestingly enough, a solution exists for both problems of overlap and no coverage. The RBF network described uses "hard" spheres. If a point falls within the sphere, then the ML node becomes active with a preset activation level. An alternative is to use "soft" spheres with a corresponding graded activation. The soft sphere has a high "density" near the center and a reduced density further out. The sphere's density can be described by a Gaussian function of the form

$$C \exp(- | r_i - r_c |^2 /(2\sigma^2))$$

where r_i is the input point, r_c is the sphere center, and sigma is a measure of the radius[12]. The ML node activation is proportional to this Gaussian. During the classification phase, all ML nodes have some amount of activation since these Gaussians have an unlimited domain. The OL nodes then sum all of the activation's from their ML nodes and the WTA selects the class with the largest overall activation. All points of feature space are included in the regions of activation of the spheres and the point-to-boundary distances are indirectly computed on the basis of the sphere activation's. This Gaussian scheme is much more intense computationally than the hard sphere approach. From a hardware implementation point of view, the Gaussian approach may be just as easy to implement if neuronal-like elements that have a graded response are used. The Gaussian spheres are incorporated into use in a similar way as the hard spheres, the main difference is that the decision to incorporate is based on requiring an activation above a non zero threshold instead of a zero threshold. These RBFs can form the basis for a fuzzy recognition system[16].

An enhancement to the Gaussian spheres is to allow the sphere to be elongated and skewed. The activation response is now of the form

$$C \exp(- (r_i - r_c)^T G(r_i - r_c))$$

where G is a positive-definite matrix that contains the sizes and skewness. A computational alternative to the skewed Gaussian spheres is to use hard hyper-rectangles. The "sphere" of influence is a hard rectangle in the feature space. The ML node must remember the center coordinates as well as the distances from the center to each face measured parallel to the appropriate feature space axis. These distances are treated in exactly the same way as the hard sphere radius for generalization and discrimination, except that now the face distances are treated individually. A limitation is that the hyper-rectangle cannot change orientation. If these faces were allowed to change orientation, then the RBF network would approach the traditional feedforward Perceptron neural network with hyperplanar boundaries. Investigations are in progress to develop learning schemes for the skewed Gaussian and hyper-rectangle networks.

37

5.2.6 RBF Application

The environment that these RBF networks are to be used is the seismological discrimination of earthquakes, quarry blasts, marine explosions, nuclear tests, etc. The seismic waveform is a set of up to 2400 samples with the beginning of the disturbance given at a specific sample. If the RBF network is to be applied to this waveform, there are too many input components to be used effectively. Then an effort must be made to extract features from this waveform. Three sets of features are moments, fractal dimensions, and wavelets[9].

The difficulty with using these features is that the waveforms are not stationary with time[19]. For example, with an earthquake, a longitudinal pressure wave is generated and propagates very fast through the crust and mantle. A transverse wave is also generated and travels at a slightly slower speed. An detector will receive the longitudinal wave (P-wave for Primary) first and the transverse wave (S-wave for Secondary) next. Other waves are generated that correspond to different oscillatory motions (e.g Love, Raleigh, etc.). Reflections and refraction's tend to split waves into components that propagate at different velocities. The seismic waveform is, therefore, a superposition of many waves.

To accommodate the temporal variance, the features were taken over a sliding window of width W samples. After the features were extracted, the window skipped to the next group of W samples, and so on. These sets of features provide the additional dimension of time. In the RBF network, the ML nodes have this extra dimension added to the sphere centers and the radius calculation. Pictorially, the feature space is like a block of wood with time along one side and the collection of RBF spheres are like wormholes in the wood.

In this application, the hard spheres are used and the learning process occurs just as before. The classification is based on the number of spheres a sequence of feature points falls in for each class. The class with the largest number gives the answer.

A problem with this scheme deals with the normalization and scaling of the data. It is necessary for the feature space to be fixed and finite and it is helpful if the feature space is the unit hypercube. To this end, the classification procedure begins with finding all feature measurements for all windows and then scaling all of the measurements for each component so that the smallest is zero and the largest is one. However, the amount for scaling and shifting differs for different waveforms and may create an unwanted variation from waveform to waveform. Resolution of this problem is still under investigation.

5.2.7 RBF Results

Waveforms of length 2400 samples were taken in windows of 80. The network consisted of 10 input nodes, 500 ML nodes and 3 OL nodes for discriminating between local events (LB), quarry blasts (QB), and regional events (R). Due to the initial lack of access to the seismic database, only 19 waveforms were obtained for this analysis. The procedure began with the feature extraction to obtain 30 10-component vectors by means of generating the first ten

38

moments from the sliding window. Then the RBF network was trained on 18 waveforms and tested on the remainder. The results were encouraging in that 70% of the waveforms were correctly classified. A deeper probing revealed that only 236 ML nodes were used out of a possible maximum of 540 nodes if each sample had to be memorized, indicating that some generalization was occurring. This classification is comparable to the backpropagation results for this data, yet the training time was about two orders of magnitude faster with the RBF network.

The second stage of the classification was to test using wavelets, since presumably the wavelet is characteristic of the mechanism that generates the seismic wave and removes the effects of propagation and reflections[9-]. Here the results were comparable to the moments in that 68% were classified correctly.

APPENDIX A   Data Base Wave Form Files from CSS

FNAME	STA	CHAN	JDATE
Febme1.w	ARU	bz	1991119
Febme16.w	ESLA	sz	1991114
Febme17.w	ESLA	sz	1991114
Febme18.w	ESLA	sz	1991135
Febme19.w	ESLA	sz	1991135
Febme43.w	GAR	bz	1990051
Febme45.w	GAR	bz	1991124
Febme47.w	GAR	bz	1991139
Febme48.w	GAR	bz	1991141
Febme49.w	GAR	bz	1991146
Febme55.w	KIV	bz	1991133
Febme56.w	KIV	bz	1991146
Febme65.w	OBN	bz	1991139
Febme66.w	OBN	bz	1991144
Febme67.w	OBN	bz	1991146
Febr0.w	GRA1	bz	1990331
Febr9.w	GRA1	bz	1991117
Febr15.w	GRA1	bz	1991127
Febr21.w	GRA1	bz	1991136
Febr46.w	WRA	sz	1990331
Febr52.w	WRA	cb	1991114
Febr58.w	WRA	cb	1991119
Febr66.w	WRA	cb	1991121
Febr72.w	WRA	cb	1991129
Febr86.w	WRA	cb	1991141
Febr99.w	WRA	cb	1991143
Febr103.w	WRA	cb	1991147
Febr109.w	WRA	cb	1991151
Febr112.w	WRA	cb	1991152
Febr115.w	WRA	cb	1991153

NOTE: All signals are 2400 samples at 20.00 samples per second.

FNAME	STA	CHAN	JDATE
Febta25.w	GRA1	bz	1991132
Febta52.w	WRA	sz	1990123
Febta69.w	WRA	sz	1990334
Febta78.w	WRA	sz	1990335
Febta81.w	WRA	sz	1990335
Febta86.w	WRA	sz	1990051
Febta97.w	WRA	sz	1990065
Febta150.w	WRA	cb	1991114
Febta177.w	WRA	cb	1991118
Febta229.w	WRA	cb	1991121
Febta309.w	WRA	cb	1991125
Febta317.w	WRA	cb	1991125
Febta408.w	WRA	cb	1991133
Febta513.w	WRA	cb	1991137
Febta542.w	WRA	cb	1991138
Febla0.w	BJT	sz	1991147
Febla5.w	GAR	bz	1991115
Febla7.w	GAR	bz	1991117
Febla8.w	GAR	bz	1991119
Febla9.w	GAR	bz	1991145
Febla11.w	GRA1	bz	1991112
Febla13.w	GRA1	bz	1991116
Febla16.w	GRA1	bz	1991122
Febla19.w	GRA1	bz	1991149
Febla20.w	HFS	sz	1991135
Febla26.w	HFS	cb	1991135
Febla73.w	WRA	cb	1991137
Febla75.w	WRA	cb	1991143
Febla76.w	WRA	cb	1991143
Febla82.w	WRA	cb	1991146

NOTE: All signals are 2400 samples at 20.00 samples per second.

FNAME	STA	CHAN	JDATE
Febqb0.w	ASAR	cb	1991123
Febqb12.w	CTA	bz	1991123
Febqb20.w	CTA	bz	1991141
Febqb33.w	KAF	sz	1990331
Febqb45.w	KAF	sz	1991114
Febqb93.w	KAF	sz	1991133
Febqb100.w	KAF	sz	1991135
Febqb114.w	KAF	sz	1991140
Febqb117.w	KAF	sz	1991140
Febqb118.w	KAF	sz	1991140
Febqb122.w	KAF	sz	1991142
Febqb147.w	KAF	sz	1991150
Febqb154.w	KAF	sz	1991154
Febqb158.w	STK	bz	1991121
Febqb180.w	WRA	cb	1991141

FNAME	STA	E-TPYE	CHAN	JDATE
2850	KAF	R	SZ	1990044
2854	KAF	R	SZ	1990044
340281	KAF	R	SZ	1990331
347028	KAF	LB	SZ	1991112
5908	KAF	R	SZ	1990065
4709	KAF	LB	SZ	1990058
423781	KAF	QB	SZ	1991152
422160	KAF	LB	SZ	1991151
418260	KAF	LB	SZ	1991149
416469	KAF	LB	SZ	1991148
386423	KAF	LB	SZ	1991133
371268	KAF	QB	SZ	1991125
360285	KAF	QB	SZ	1991120
356908	KAF	QB	SZ	1991118
347142	KAF	QB	SZ	1991113
355627	KAF	R	SZ	1991117
379845	KAF	QB	SZ	1991129
381583	KAF	QB	SZ	1991130
351941	KAF	R	SZ	1991115

NOTE: All signals are 2400 samples at 20.00 samples per second.

GSETT-Subset1 Station Names and Locations

STA	STATION NAME	LATITUDE	LONGITUDE
ARU	ARTI - SVERDLOVSK, OBLAST	56.4000	58.6000
ASAR	ALICE SPRINGS ARRAY - NORTH TERRITORY, AUSTRALIA	23.7040	133.9620
BJT	BAIJIATUAN - BAIJIATUAN, CHINA	40.0403	116.1750
CTA	CHARTERS TOWERS - QUEENSLAND, AUSTRALIA	20.0880	146.2540
ESLA	SONSECA ARRAY STATION - SPAIN	39.6700	-3.9600
GAR	GARM - GARM, USSR	39.0000	70.3000
GRA1	GRAFENBERG ARRAY - BOYERN, GERMANY	49.6920	11.2220
HFS	HAGFORS ARRAY - SWEDEN	60.1335	13.6836
KAF	KANGASNIEMI - FINLAND	62.1127	26.3062
KIV	KISLOVODSK - WESTERN CAUCASUS USSR	43.9500	42.6833
OBN	OBNINSK - OBNINSK, USSR	55.1167	36.5667
STK	STEPHENS CREEK - NEW SOUTH WALES, AUSTRALIA	31.8820	141.5920
WRA	WARRAMUNGA ARRAY - NORTHERN TERRITORY, AUSTRALIA	-19.7657	134.3891

APPENDIX B   Backpropagation Neural Network

Back Propagation is an iterative gradient descent method which seeks to minimize the mean-square error. This - among other things - means that the updating rule is the so-called delta rule. In the delta rule, we have new weight = old weight + delta*error where delta is a learning rate - often taken to be 0.1. Back propagation was among the first methods which allowed the training of hidden neurons in a multi-layer neural network. It had always been possible to find the error in an output neuron. It was simply defined to be the absolute value of the difference between the actual output and the desired output. For a hidden neuron - say n[i]- the error was a bit more nebulous. In back propagation, error in a hidden node was defined as follows:

Let e[i] represent the error in a hidden neuron n[i] and suppose that n[i] is connected to neurons n[j], then e[i] is defined by the equation

$$e[i] = w[i,j]*e[j]$$

where e[j] is the error in neuron j and w[i,j] is the weight from neuron i to neuron j. From this definition of error for hidden nodes and with gradient descent as a training method, we get the method of back propagation. Back propagation also uses "Squashing" or "Sigmoidal" functions to insure that all neurons (hidden or otherwise) produce outputs in [0,1]. The most commonly used function is given by:

$$f(x)=1/(1+exp(-(x-k))$$

where k is some constant.

Training A BackProp Neural Net

The steps involved in training a back propagation neural network are:

1. Weight initialization.

Typically all weights are set to small random values in [0,1]. This method is employed for lack of something better to do rather than some deep mathematical reason.

2. Presenting Input and Desired Output

BackProp is a supervised neural network, so the desired output ispresented each time an input vector is presented tothe network. The input vector may be thought of asa continuous valued vector. The output vector is generally a binary vector i.e. each entry is 0 or 1. The output is the weighted sum of the input values to a neuron times the corresponding weights. Once this value has been calculated, it is passed through the "Squashing" or sigmoid function to give the final value for the output.

44

3. . Adapting the Weights

Starting at the output nodes and working backwards to the input nodes, the weights are adjusted by the delta rule. The formulas are slightly different for weights on connections to output neurons than the ones for those in hidden layers. This is due to the way error is calculated.

4. Iterating

The process is repeated by going back to step ii. Training often stops when
   a. things look hopeless
   b. the net has learned the training set
   c. a set number of iterations have been done
   d. an acceptable percentage of the training set has been learned.

Adjusting Parameters

The parameters which are most often adjusted in BackProp are:

   a. delta or the training rate
   b. the number of layers
   c. the number of neurons in each layer
   d. the constant k in the sigmoidal function.

No good rules exist for choosing or adjusting any of the parameters given above.

Back Propagation in Detail

The diagram given below is intended to serve as a guide for an n-layered backprop neural network. We shall make the assumptions that:

   1.   layer 1 is the input layer
   2.   layer n is the output layer
   3.   w[i,j,k] is the weight from neuron j in layer i to neuron k in layer i+1.
   4.   no_in[i] is the number of neurons in layer i
   5.   e[i,j] is the error in neuron j in layer i
   6.   out[i,j] is the output of neuron j in layer i

A Procedure to Initialize Weights

Note that an n-layered network will have n-1 sets of weights.

    procedure initialize_weights (w,no_in,n)

```
begin
  for i = 1 to n-1
    for j = 1 to no_in[i]
      for k = 1 to no_in[i+1]
        w[i,j,k] = random;
      end
    end
  end
end  { end procedure }
```

A Procedure to Compute the Outputs

Note that only layers 2-n have output. The output of layer 1 is the input vector which we shall call Y. Note that Y should have no_in[1] components.

```
procedure compute_outputs (w,n,no_in)

begin
  { Transfer the input vector to layer 1 }
  for j = 1 to no_in[1]
    out[1,j] = Y[j];
  end;

  for i = 2 to n do
    for j = 1 to no_in[i]
      out[i,j]=0;
      for k:=1 to no_in[i-1]
        out[i,j]=out[i,j] +w[i-1,j,k]*out[i-1,k]
      end { end k }
      out[i,j] = 1/(1+exp(-out[i,j])
    end ; { end j loop }
  end;

end; { end compute outputs }
```

A Procedure to Update the Weights

```
Procedure update_weights(n,w,delta,desired)

begin
  for i = n-1 downto 1 do
    for j = 1 to no_in[i]
```

46

```
      if i = n-1 then  { a weight to output layer }
        e[i,j] = output[i,j]*(1-output[i,j])*
               (desired[j]-output[i,j])
      else begin
        e[i,j] = 0;
        for k = 1 to no_in[i+1]
          e[i,j]=e[i,j]+w[i,j,k]*e[i+1,k]
        end { end k }
        e[i,j]=e[i,j]*out[i,j]*(1-out[i,j])
      end { end else }

   end; { end i loop }

end; { end update weights }
```

APPENDIX C   Unsupervised Kohonen Networks

Suppose that a set of data consists of M points which fall into one of N classes. This number - N - may or may not be known.

Example - A hospital takes data from each of its 1000 patients and records the results in a patient vector. Then M = 1000. It is desired to use the patient vectors to determine which patients have the same disease. Thus N - the number of classes is the same as the number of diseases and N may or may not be known.

In Unsupervised Kohonen neural networks, a set of neurons is trained to arrange themselves at or near the centers of the classes. When training is over, this set of neurons (Kohonen called them codebook vectors) is able to give an idea of N or the number of distinct classes in the data classify an unknown input vector by nearest neighbor where "near" may mean Euclidean distance or some other measureof distance.

The network is not able to tell what the individual classes are. Thus in the hospital example, a Kohonen Neural Network could place all individuals with the same disease in the same class but it could not assign a name to the disease.

This is very similar to Cluster Analysis in statistics and networks such as the k-means neural network. The literature generally reports that Kohonen networks are very slow to train (by design) are good pattern recognizers are noise tolerant.

Problems with Kohonen networks include how to know when to stop training (choice of training parameters), how to initialize the codebook vectors and the appropriate number of codebook vectors

Training an Unsupervised Kohonen Net

Note: Neuron and codebook vector are used interchangeably in the following discussion.

Let N be "comfortably large" and define an array of N vectors with C components. In the hospital example we might make the following analysis:

Suppose we are reasonably sure that there are at most 25 diseases among the 1000 patients. Suppose that we took 10 measurements from each patient e.g. temperature, blood pressure, blood count, etc. We might decide to begin with 50 neurons with 10 components. This forms our Codebook.

If we choose too few neurons, the data cannot be reasonably covered. If we choose too many newurons, some of the diseases may subdivide into subclasses which are not distinguishable even to a trained observer. There are no good rules of thumb to follow concerning the number of neurons vs. the population size. One generally has to experiment to find a good number.

Initializing the Codebook

As stated above, initializing the codebook is a difficult problem. Some suggested ways are given below:

a. Assign each codebook vector a random value.
```
for i = 1 to N   { for each codebook vector}
  for j = 1 to C
    codebook[i,j] = random
```

b. Assign each codebook vector the same constant value
```
for i = 1 to N   { for each codebook vector}
  for j = 1 to C
    codebook[i,j] = K { K is a constant }
```

c. If the range of values of each component is known, assign component j a random value in the range of that component. Say that component i varies from a maximum value of Max[i] to a minimum value of Min[i]. We could code

```
for i = 1 to N   { for each codebook vector}
  for j = 1 to C
    codebook[i,j] = random*(Max[j]-Min[j]) + Min[j]
```

d. Similar to c is the approach that assigns component j the average value of that component for the entire data set. This is intractible for large data sets. If however this value is known, we may code

```
for i = 1 to N   { for each codebook vector}
  for j = 1 to C
    codebook[i,j] = avg[j]  { avg[j] is the average of
                             all of the jth components}
```

e. Something else that may make sense.

Training the Network

The essential part of Kohonen training is summarized as follows:

0. Let lambda be a training rate (Kohonen has suggested 0.2) and let Maxiter denote the number of training iterations you wish to perform.

49

1. Let X be an input vector to the training procedure.

2. Compute the distance from X to each of the codebook vectors. Distance could mean Euclidean distance in N space or it could mean the cosine of the angle between X and each codebook vector.

3. Let k' be the closest codebook vector. k' is often referred to as the winning or firing neuron.

4. Update code vector k' by the following formula
   for j = 1 to C
      codebook[k',j] = codebook[k',j] +
         lambda*(X[j]-codebook[k',j]

5. Decrement lambda

6. If you have reached Maxiter or lambda has reached 0, terminate training. Otherwise repeat steps 1-6.

Variations of Kohonen Training

As stated earlier, the initialization of the Kohonen neurons is a difficult problem particularly in the absence of information about the data. This often leads to the problem of too few neurons to cover the space - or in the other extreme too many neurons which break the data down into meaningless classes. To overcome this problem some suggested solutions are given below.

1. For the first "several" passes through the data file, update every codebook neuron. This has the effect of pulling all of the neurons to "where the data is".

2. Keep a record of how many times a neuron has fired. If it does not fire "in a long while", force it to fire. This simply means to update the winning neuron and the idle neuron by the formula given in step 4 above.

Complete Pseudocode for Unsupervised Kohonen

```
CONST
  Maxiter = -----;
  lambda  = -----;
  C= -----;  { Number of components in vectors }
  N= -----;  { Number of codebook vectors }
```

```
Procedure initialize_codebook_neurons(codebook,C,N)

        begin
          for i = 1 to N
            for j = 1 to C
              codebook[i,j] =   a. random value
                            b. average value of jth component
                            c. constant value
                            d. random value in the range of
                               jth component

        end;


Procedure Compute_distances (Codebook,X,C,N,k_prime )

        begin
        FOR EUCLIDEAN DISTANCE CODE THE FOLLOWING AND OMIT BELOW
          for i = 1 to n;
            Dist[i]=0;
            for j = 1 to c
              dist[i]=dist[i]+sqr(X[j]-codebook[i,j]);
          end;
          k_prime = 1;
          for i = 2 to N
            if Dist[i] < dist[k_prime];
              k_prime = i;
          end;


        FOR MAXIMUM DOT PRODUCT CODE THE FOLLOWING AND OMIT ABOVE
          for i = 1 to n
            Dist[i]=0;
            norm_x=0;
            norm_neuron = 0;
            for j = 1 to C
              dist[i]=dist[i]+(X[j]*codebook[i,j])
              norm_x = norm_x + sqr(X[j]);
              norm_neuron=norm_neuron+sqr(codebook[i,j];
            end;
            dist[i]= dist[i]/(sqrt(norm_x)*sqrt(norm_neuron);
          end;

          k_prime = 1;
          for i = 2 to N
            if Dist[i] > dist[k_prime]
```

```
          k_prime = i
    end;

  NOTE THAT k_prime is the index of the winning neuron

  end; { end compute_distances }


Procedure update_neuron(codebook,X,k)
  begin
    for j = 1 to C
      codebook[k,j]=codebook[k,j]+lambda*(codebook[k,j]- X[j]);
    end;
  end;

Begin { Begin main }

    initialize_codebook_neurons(codebook,C,N);
    trained = false;
    lambda0=lambda;
    iterations=0;
    while trained = false
      read input vector X
      compute_distances(codebook,X,C,N,k_prime);
      update_neuron(codebook,X,k_prime);
      if iterations < (you pick it)
        begin
          for i = 1 to N
            update_neuron(codebook,X,i);
        iter=iter +1;
        lambda=lambda-lambda0/Maxiter;
      end {while}

  end;
```

52

APPENDIX D   Supervised Kohonen Networks

Training the Network

The essential parts of Supervised Kohonen training are summarized as follows. Those steps marked with * are identical to the corresponding step in unsupervised learning. The reader may observe that if the neural network is correct, the codebook vector is rotated toward the input vector. If the network is incorrect, the codebook vector is rotated away from the input vector.

*0. Let lambda be a training rate (Kohonen has suggested 0.2)
and let Maxiter denote the number of training iterations you
wish to perform.

1. Let X be an input vector to the training procedure with a
known classification, say x_class.

*2. Compute the distance from X to each of the codebook vectors.
Distance could mean Euclidean distance in N space or it
could mean the cosine of the angle between X and each
codebook vector.

*3. Let k' be the closest codebook vector. k' is often referred
to as the winning or firing neuron.

4. Update code vector k' by the following formula
   if the classification of X as belonging to the class
   represented by the vector k' is correct then
   for j = 1 to C
     codebook[k',j] = codebook[k',j] +
         lambda*(X[j]-codebook[k',j]
   else
   for j = 1 to C
     codebook[k',j] = codebook[k',j] -
         lambda*(X[j]-codebook[k',j]

* 5. Decrement lambda

* 6. If you have reached Maxiter or lambda has reached 0,
terminate training. Otherwise repeat steps 1-6.

Complete Pseudocode for Supervised Kohonen

```
CONST
  Maxiter = ——;
  lambda  = ——;
  C= ——;  { Number of components in vectors }
  N= ——;  { Number of codebook vectors }

Procedure initialize_codebook_neurons(codebook,C,N)

 begin
  for i = 1 to N
   for j = 1 to C
     codebook[i,j] =  a. random value
                      b. average value of jth component
                      c. constant value
                      d. random value in the range of
                         jth component
   for i = 1 to N
    neuron_id[i] = class to be represented by codebook
     vector i

 end;


Procedure Compute_distances (Codebook,X,C,N,k_prime )

  begin
  FOR EUCLIDEAN DISTANCE CODE THE FOLLOWING AND OMIT BELOW
   for i = 1 to n;
    Dist[i]=0;
    for j = 1 to c
      dist[i]=dist[i]+sqr(X[j]-codebook[i,j]);
   end;

   k_prime = 1;

for i = 2 to N
    if Dist[i] < dist[k_prime];
      k_prime = i;
    end;

FOR MAXIMUM DOT PRODUCT CODE THE FOLLOWING AND OMIT ABOVE

   for i = 1 to n
```

```
  Dist[i]=0;
  norm_x=0;
  norm_neuron = 0;
  for j = 1 to C
    dist[i]=dist[i]+(X[j]*codebook[i,j])
    norm_x = norm_x + sqr(X[j]);
    norm_neuron=norm_neuron+sqr(codebook[i,j];
  end;
    dist[i]= dist[i]/(sqrt(norm_x)*sqrt(norm_neuron);
  end;

k_prime = 1;
for i = 2 to N
  if Dist[i] > dist[k_prime]
    k_prime = i
end;

NOTE THAT k_prime is the index of the winning neuron

end; { end compute_distances }


Procedure update_neuron(correct,codebook,X,k)
  begin
  if correct = true then
    for j = 1 to C
      codebook[k,j]=codebook[k,j]+lambda*(codebook[k,j]- X[j]);
    end;
   else
    for j = 1 to C
      codebook[k,j]=codebook[k,j]-lambda*(codebook[k,j]- X[j]);
    end;

  end;


Begin { Begin main }

    initialize_codebook_neurons(codebook,C,N);
    trained = false;
    lambda0=lambda;
    iterations=0;
    while trained = false
      read input vector X and its class - say x_class
      compute_distances(codebook,X,C,N,k_prime);
```

```
      if neuron_id[kprime] = x_class then
         correct = true
      else
         correct = true;
      update_neuron(correct,codebook,X,k_prime);
      if iterations < (you pick it)
        begin
          for i = 1 to N
            update_neuron(codebook,X,i);
        iter=iter +1;
        lambda=lambda-lambda0/Maxiter;
    end {while}

end;
```

REFERENCES

[1] ADA Methodologies: Concepts and Requirements, U.S. Department of Defense, December 1982.

[2] "Artificial Neural Networks for Seismic Data Interpretation," Lexington, Massachusetts, MIT Lincoln Laboratory, Semi-Annual Technical Summary, November 30, 1990.

[3] "Artificial Neural Networks for Seismic Data Interpretation," Lexington, Massachusetts, MIT Lincoln Laboratory, Semi-Annual Technical Summary, June 30, 1992.

[4] Bache, Thomas, C., et. al., "The Intelligent Monitoring System," Bulletin of Seismological Society of America, Volume 80, Number 6, December 1990, pp. 1833-1851.

[5] Bratt, Steve, "Data Management and Access at the Center for Seismic Studies", Center for Seismic Studies, November 1991.

[6] Cohen, Norman H., ADA as a Second Language, McGraw-Hill, 1986.

[7] Collard, Philippe, and Andre Goforth, "Knowledge Based Systems and ADA: An Overview of the Issues," ADA Letters, November/December 1988, Volume VIII, Number 6, pp. 72-81.

[8] Dahlmann, O. and H. Israelson, Monitoring Underground Nuclear Explosions, Amsterdam (Elsevier), 1977.

[9] Davari, A., Seismic Data Processing via Homomorphic Filtering, SSST 1993, Alabama.

[10] Dowla, Farid U., et. all, "Seismic Discrimination with Artificial Neural Networks: Preliminary Results with Regional Spectral Data", Bulletin of the Seismological Society of America, vol. 80, no. 5, pp. 1346-1373, October 1990.

[11] Dysart, Paul S., and Jay J. Pulli, "Regional Seismic Event Classification at the Noress Array: Seismological Measurements and the use of Trained Neural Networks", Bulletin of the Seismological Society of America, vol. 80, no. 6, pp. 1910-1933, December 1990.

[12] Hanson, S.J., and Gluck, M. A., Spherical Units as Dynamic Consequential Regions:Implications for Attention, Competition, and Categorization, in *Advances in Neural Information Processing Systems*, Vol.3, pp.656-664, 1991.

[13] Hemmendinger, David, "Specifying ADA Server Tasks with Executable Formal Grammars," IEEE Transactions on Software Engineering, Volume 16, Number 7, July 1990, pp. 741-754.

[14] Hedlin, A.H., J.B. Minster and J.A. Orcutt, "An Automatic Means to Discriminate Between Earthquakes and Quarry Blasts", Bulletin of the Seismological Society of America, vol. 80, no. 6, pp. 12143-2160, December 1990.

[15] Hedlin, A.H., J.B. Minster and J.A. Orcutt,"The time-frequency characteristic of quarry blasts and calibration explosions recoreded in Kazakhstan, USSR," Geophysics Journal International 89,109-121.

[16] Jokinen, Petri A., On the Relation Between Radial Basis Function Networks and Fuzzy Systems, *International Joint Conference on Neural Networks*, Vol.1, pp.221-225, Baltimore, 1992.

[17] Joswig, Manfred, "Pattern Recognition for Earthquake Detection," Bulletin of the Seismological Society of America, Volume 80, Number 1, February 1990, pp. 170-186.

[18] Jurkevics, Andy, "Polarization Analysis of Three-Component Array Data," Bulletin of the Seismology Society of America, Volume 78, Number 5, October 1988, pp. 1725-1743.

[19] Kulhanek, Ota, ANATOMY OF SEISMOGRAMS, Elsevier, 1990.

[20] McClellan, James H., Two-Dimensional Spectrum Analysis in Sonic Logging," IEEE ASSP Magazine, July 1986, pp. 12-18.

[21] Mendel, Jerry M., "Some Modeling Problems in Reflection Seismology," IEEE ASSP Magazine, April 1986, pp. 4-17.

[22] Moody, J. and Darken, C., Fast Learning in Networks of Locally-Tuned Processing Units, *Neural Computation*, Vol.1, pp.281-294, 1989.

[23] Mykkeltveit, Svein, et al, "Application of Regional Arrays in Seismic Verification Research," Bulletin of Seismological Society of America, Vol. 80, Number 6, December 1990, pp.1777-1800.

[24] Mykkeltveit, Svein, and H. Bungum, "Processing of Regional Seismic Events Using Data From Small-Aperture Arrays," Bulletin of Seismological Society of America, Volume 74, Number 6, December 1984, pp. 2313-2333.

[25] Papoulis, Athanusios, "Probability," Random Variables, and Stochastic Processes, 3rd Edition, McGraw-Hill, 1991.

[26] Reberto, Vito, and Claudio Chiaruttini, "Seismic Signal Understanding: A Knowledge-Based System", IEEE Transactions on Signal Processing, vol. 40, no. 7, July 1992, pp. 1787-1806.

[27] Roberts, R. G., A. Christoffersson, and F. Cassidy; "Real Time Event Detection, Phase Identification, and Source Location Estimation Using Single Station Three-Component Seismic Data, Geophysical Journal, Number 97, 1989, pp. 471-480.

[28] Sammet, Jean E., "Why ADA is Not Just Another Programming Language," Communications of the ACM, Volume 29, Number 8, August 1986, pp. 722-732.

[29] Schach, Stephen R., Software Engineering, Aksen Associates, 1990.

[30] Shumate, Ken, Understanding ADA, Harper and Row, 1984.

[31] Turbo Prolog Owner's Handbook, Borland International, 1986.

[32] Wiener, Richard, and Richard Sincovec, Programming in ADA, Wiley, 1983.

[33] Wuster. Jan, "Discrimination of Chemical Explosions and Earthquakes in Central Europe - A Case Study", DARPA Annual Report, 1992, under grant AFOSR-90-0189.

Prof. Thomas Ahrens
Seismological Lab, 252-21
Division of Geological & Planetary Sciences
California Institute of Technology
Pasadena, CA 91125

Prof. Keiiti Aki
Center for Earth Sciences
University of Southern California
University Park
Los Angeles, CA 90089-0741

Prof. Shelton Alexander
Geosciences Department
403 Deike Building
The Pennsylvania State University
University Park, PA 16802

Dr. Ralph Alewine, III
DARPA/NMRO
3701 North Fairfax Drive
Arlington, VA 22203-1714

Prof. Charles B. Archambeau
CIRES
University of Colorado
Boulder, CO 80309

Dr. Thomas C. Bache, Jr.
Science Applications Int'l Corp.
10260 Campus Point Drive
San Diego, CA 92121 (2 copies)

Prof. Muawia Barazangi
Institute for the Study of the Continent
Cornell University
Ithaca, NY 14853

Dr. Jeff Barker
Department of Geological Sciences
State University of New York
 at Binghamton
Vestal, NY 13901

Dr. Douglas R. Baumgardt
ENSCO, Inc
5400 Port Royal Road
Springfield, VA 22151-2388

Dr. Susan Beck
Department of Geosciences
Building #77
University of Arizona
Tuscon, AZ 85721

Dr. T.J. Bennett
S-CUBED
A Division of Maxwell Laboratories
11800 Sunrise Valley Drive, Suite 1212
Reston, VA 22091

Dr. Robert Blandford
AFTAC/TT, Center for Seismic Studies
1300 North 17th Street
Suite 1450
Arlington, VA 22209-2308

Dr. Stephen Bratt
Center for Seismic Studies
1300 North 17th Street
Suite 1450
Arlington, VA 22209-2308

Dr. Lawrence Burdick
IGPP, A-025
Scripps Institute of Oceanography
University of California, San Diego
La Jolla, CA 92093

Dr. Robert Burridge
Schlumberger-Doll Research Center
Old Quarry Road
Ridgefield, CT 06877

Dr. Jerry Carter
Center for Seismic Studies
1300 North 17th Street
Suite 1450
Arlington, VA 22209-2308

Dr. Eric Chael
Division 9241
Sandia Laboratory
Albuquerque, NM 87185

Dr. Martin Chapman
Department of Geological Sciences
Virginia Polytechnical Institute
21044 Derring Hall
Blacksburg, VA 24061

Prof. Vernon F. Cormier
Department of Geology & Geophysics
U-45, Room 207
University of Connecticut
Storrs, CT 06268

Prof. Steven Day
Department of Geological Sciences
San Diego State University
San Diego, CA 92182

1

Marvin Denny
U.S. Department of Energy
Office of Arms Control
Washington, DC 20585

Dr. Zoltan Der
ENSCO, Inc.
5400 Port Royal Road
Springfield, VA 22151-2388

Prof. Adam Dziewonski
Hoffman Laboratory, Harvard University
Dept. of Earth Atmos. & Planetary Sciences
20 Oxford Street
Cambridge, MA 02138

Prof. John Ebel
Department of Geology & Geophysics
Boston College
Chestnut Hill, MA 02167

Eric Fielding
SNEE Hall
INSTOC
Cornell University
Ithaca, NY 14853

Dr. Mark D. Fisk
Mission Research Corporation
735 State Street
P.O. Drawer 719
Santa Barbara, CA 93102

Prof Stanley Flatte
Applied Sciences Building
University of California, Santa Cruz
Santa Cruz, CA 95064

Dr. John Foley
NER-Geo Sciences
1100 Crown Colony Drive
Quincy, MA 02169

Prof. Donald Forsyth
Department of Geological Sciences
Brown University
Providence, RI 02912

Dr. Art Frankel
U.S. Geological Survey
922 National Center
Reston, VA 22092

Dr. Cliff Frolich
Institute of Geophysics
8701 North Mopac
Austin, TX 78759

Dr. Holly Given
IGPP, A-025
Scripps Institute of Oceanography
University of California, San Diego
La Jolla, CA 92093

Dr. Jeffrey W. Given
SAIC
10260 Campus Point Drive
San Diego, CA 92121

Dr. Dale Glover
Defense Intelligence Agency
ATTN: ODT-1B
Washington, DC 20301

Dr. Indra Gupta
Teledyne Geotech
314 Montgomery Street
Alexanderia, VA 22314

Dan N. Hagedon
Pacific Northwest Laboratories
Battelle Boulevard
Richland, WA 99352

Dr. James Hannon
Lawrence Livermore National Laboratory
P.O. Box 808
L-205
Livermore, CA 94550

Dr. Roger Hansen
HQ AFTAC/TTR
130 South Highway A1A
Patrick AFB, FL 32925-3002

Prof. David G. Harkrider
Seismological Laboratory
Division of Geological & Planetary Sciences
California Institute of Technology
Pasadena, CA 91125

Prof. Danny Harvey
CIRES
University of Colorado
Boulder, CO 80309

2

Prof. Donald V. Helmberger
Seismological Laboratory
Division of Geological & Planetary Sciences
California Institute of Technology
Pasadena, CA 91125

Prof. Eugene Herrin
Institute for the Study of Earth and Man
Geophysical Laboratory
Southern Methodist University
Dallas, TX 75275

Prof. Robert B. Herrmann
Department of Earth & Atmospheric Sciences
St. Louis University
St. Louis, MO 63156

Prof. Lane R. Johnson
Seismographic Station
University of California
Berkeley, CA 94720

Prof. Thomas H. Jordan
Department of Earth, Atmospheric &
  Planetary Sciences
Massachusetts Institute of Technology
Cambridge, MA 02139

Prof. Alan Kafka
Department of Geology & Geophysics
Boston College
Chestnut Hill, MA 02167

Robert C. Kemerait
ENSCO, Inc.
445 Pineda Court
Melbourne, FL 32940

Dr. Karl Koch
Institute for the Study of Earth and Man
Geophysical Laboratory
Southern Methodist University
Dallas, Tx 75275

Dr. Max Koontz
U.S. Dept. of Energy/DP 5
Forrestal Building
1000 Independence Avenue
Washington, DC 20585

Dr. Richard LaCoss
MIT Lincoln Laboratory, M-200B
P.O. Box 73
Lexington, MA 02173-0073

Dr. Fred K. Lamb
University of Illinois at Urbana-Champaign
Department of Physics
1110 West Green Street
Urbana, IL 61801

Prof. Charles A. Langston
Geosciences Department
403 Deike Building
The Pennsylvania State University
University Park, PA 16802

Jim Lawson, Chief Geophysicist
Oklahoma Geological Survey
Oklahoma Geophysical Observatory
P.O. Box 8
Leonard, OK 74043-0008

Prof. Thorne Lay
Institute of Tectonics
Earth Science Board
University of California, Santa Cruz
Santa Cruz, CA 95064

Dr. William Leith
U.S. Geological Survey
Mail Stop 928
Reston, VA 22092

Mr. James F. Lewkowicz
Phillips Laboratory/GPEH
29 Randolph Road
Hanscom AFB, MA 01731-3010(2 copies)

Mr. Alfred Lieberman
ACDA/VI-OA State Department Building
Room 5726
320-21st Street, NW
Washington, DC 20451

Prof. L. Timothy Long
School of Geophysical Sciences
Georgia Institute of Technology
Atlanta, GA 30332

Dr. Randolph Martin, III
New England Research, Inc.
76 Olcott Drive
White River Junction, VT 05001

Dr. Robert Masse
Denver Federal Building
Box 25046, Mail Stop 967
Denver, CO 80225

Dr. Gary McCartor
Department of Physics
Southern Methodist University
Dallas, TX 75275

Dr. Bao Nguyen
HQ AFTAC/TTR
130 South Highway A1A
Patrick AFB, FL 32925-3002

Prof. Thomas V. McEvilly
Seismographic Station
University of California
Berkeley, CA 94720

Prof. John A. Orcutt
IGPP, A-025
Scripps Institute of Oceanography
University of California, San Diego
La Jolla, CA 92093

Dr. Art McGarr
U.S. Geological Survey
Mail Stop 977
U.S. Geological Survey
Menlo Park, CA 94025

Prof. Jeffrey Park
Kline Geology Laboratory
P.O. Box 6666
New Haven, CT 06511-8130

Dr. Keith L. McLaughlin
S-CUBED
A Division of Maxwell Laboratory
P.O. Box 1620
La Jolla, CA 92038-1620

Dr. Howard Patton
Lawrence Livermore National Laboratory
L-025
P.O. Box 808
Livermore, CA 94550

Stephen Miller & Dr. Alexander Florence
SRI International
333 Ravenswood Avenue
Box AF 116
Menlo Park, CA 94025-3493

Dr. Frank Pilotte
HQ AFTAC/TT
130 South Highway A1A
Patrick AFB, FL 32925-3002

Prof. Bernard Minster
IGPP, A-025
Scripps Institute of Oceanography
University of California, San Diego
La Jolla, CA 92093

Dr. Jay J. Pulli
Radix Systems, Inc.
201 Perry Parkway
Gaithersburg, MD 20877

Prof. Brian J. Mitchell
Department of Earth & Atmospheric Sciences
St. Louis University
St. Louis, MO 63156

Dr. Robert Reinke
ATTN: FCTVTD
Field Command
Defense Nuclear Agency
Kirtland AFB, NM 87115

Mr. Jack Murphy
S-CUBED
A Division of Maxwell Laboratory
11800 Sunrise Valley Drive, Suite 1212
Reston, VA 22091 (2 Copies)

Prof. Paul G. Richards
Lamont-Doherty Geological Observatory
 of Columbia University
Palisades, NY 10964

Dr. Keith K. Nakanishi
Lawrence Livermore National Laboratory
L-025
P.O. Box 808
Livermore, CA 94550

Mr. Wilmer Rivers
Teledyne Geotech
314 Montgomery Street
Alexandria, VA 22314

Dr. Carl Newton
Los Alamos National Laboratory
P.O. Box 1663
Mail Stop C335, Group ESS-3
Los Alamos, NM 87545

Dr. George Rothe
HQ AFTAC/TTR
130 South Highway A1A
Patrick AFB, FL 32925-3002

4

Dr. Alan S. Ryall, Jr.
DARPA/NMRO
3701 North Fairfax Drive
Arlington, VA 22209-1714

Dr. Richard Sailor
TASC, Inc.
55 Walkers Brook Drive
Reading, MA 01867

Prof. Charles G. Sammis
Center for Earth Sciences
University of Southern California
University Park
Los Angeles, CA 90089-0741

Prof. Christopher H. Scholz
Lamont-Doherty Geological Observatory
of Columbia University
Palisades, NY 10964

Dr. Susan Schwartz
Institute of Tectonics
1156 High Street
Santa Cruz, CA 95064

Secretary of the Air Force
(SAFRD)
Washington, DC 20330

Office of the Secretary of Defense
DDR&E
Washington, DC 20330

Thomas J. Sereno, Jr.
Science Application Int'l Corp.
10260 Campus Point Drive
San Diego, CA 92121

Dr. Michael Shore
Defense Nuclear Agency/SPSS
6801 Telegraph Road
Alexandria, VA 22310

Dr. Robert Shumway
University of California Davis
Division of Statistics
Davis, CA 95616

Dr. Matthew Sibol
Virginia Tech
Seismological Observatory
4044 Derring Hall
Blacksburg, VA 24061-0420

Prof. David G. Simpson
IRIS, Inc.
1616 North Fort Myer Drive
Suite 1050
Arlington, VA 22209

Donald L. Springer
Lawrence Livermore National Laboratory
L-025
P.O. Box 808
Livermore, CA 94550

Dr. Jeffrey Stevens
S-CUBED
A Division of Maxwell Laboratory
P.O. Box 1620
La Jolla, CA 92038-1620

Lt. Col. Jim Stobie
ATTN: AFOSR/NL
110 Duncan Avenue
Bolling AFB
Washington, DC 20332-0001

Prof. Brian Stump
Institute for the Study of Earth & Man
Geophysical Laboratory
Southern Methodist University
Dallas, TX 75275

Prof. Jeremiah Sullivan
University of Illinois at Urbana-Champaign
Department of Physics
1110 West Green Street
Urbana, IL 61801

Prof. L. Sykes
Lamont-Doherty Geological Observatory
of Columbia University
Palisades, NY 10964

Dr. David Taylor
ENSCO, Inc.
445 Pineda Court
Melbourne, FL 32940

Dr. Steven R. Taylor
Los Alamos National Laboratory
P.O. Box 1663
Mail Stop C335
Los Alamos, NM 87545

Prof. Clifford Thurber
University of Wisconsin-Madison
Department of Geology & Geophysics
1215 West Dayton Street
Madison, WS 53706

DARPA/PM
3701 North Fairfax Drive
Arlington, VA 22203-1714

Prof. M. Nafi Toksoz
Earth Resources Lab
Massachusetts Institute of Technology
42 Carleton Street
Cambridge, MA 02142

DARPA/RMO/RETRIEVAL
3701 North Fairfax Drive
Arlington, VA 22203-1714

Dr. Larry Turnbull
CIA-OSWR/NED
Washington, DC 20505

DARPA/RMO/SECURITY OFFICE
3701 North Fairfax Drive
Arlington, VA 22203-1714

Dr. Gregory van der Vink
IRIS, Inc.
1616 North Fort Myer Drive
Suite 1050
Arlington, VA 22209

HQ DNA
ATTN: Technical Library
Washington, DC 20305

Dr. Karl Veith
EG&G
5211 Auth Road
Suite 240
Suitland, MD 20746

Defense Intelligence Agency
Directorate for Scientific & Technical Intelligence
ATTN: DTIB
Washington, DC 20340-6158

Prof. Terry C. Wallace
Department of Geosciences
Building #77
University of Arizona
Tuscon, AZ 85721

Defense Technical Information Center
Cameron Station
Alexandria, VA 22314 (2 Copies)

Dr. Thomas Weaver
Los Alamos National Laboratory
P.O. Box 1663
Mail Stop C335
Los Alamos, NM 87545

TACTEC
Battelle Memorial Institute
505 King Avenue
Columbus, OH 43201 (Final Report)

Dr. William Wortman
Mission Research Corporation
8560 Cinderbed Road
Suite 700
Newington, VA 22122

Phillips Laboratory
ATTN: XPG
29 Randolph Road
Hanscom AFB, MA 01731-3010

Prof. Francis T. Wu
Department of Geological Sciences
State University of New York
 at Binghamton
Vestal, NY 13901

Phillips Laboratory
ATTN: GPE
29 Randolph Road
Hanscom AFB, MA 01731-3010

AFTAC/CA
(STINFO)
Patrick AFB, FL 32925-6001

Phillips Laboratory
ATTN: TSML
5 Wright Street
Hanscom AFB, MA 01731-3004

Phillips Laboratory
ATTN: PL/SUL
3550 Aberdeen Ave SE
Kirtland, NM 87117-5776 (2 copies)

Dr. Svein Mykkeltveit
NTNT/NORSAR
P.O. Box 51
N-2007 Kjeller, NORWAY (3 Copies)

Dr. Michel Bouchon
I.R.I.G.M.-B.P. 68
38402 St. Martin D'Heres
Cedex, FRANCE

Prof. Keith Priestley
University of Cambridge
Bullard Labs, Dept. of Earth Sciences
Madingley Rise, Madingley Road
Cambridge CB3 OEZ, ENGLAND

Dr. Michel Campillo
Observatoire de Grenoble
I.R.I.G.M.-B.P. 53
38041 Grenoble, FRANCE

Dr. Jorg Schlittenhardt
Federal Institute for Geosciences & Nat'l Res.
Postfach 510153
D-3000 Hannover 51, GERMANY

Dr. Kin Yip Chun
Geophysics Division
Physics Department
University of Toronto
Ontario, CANADA

Dr. Johannes Schweitzer
Institute of Geophysics
Ruhr University/Bochum
P.O. Box 1102148
4360 Bochum 1, GERMANY

Prof. Hans-Peter Harjes
Institute for Geophysic
Ruhr University/Bochum
P.O. Box 102148
4630 Bochum 1, GERMANY

Trust & Verify
VERTIC
8 John Adam Street
London WC2N 6EZ, ENGLAND

Prof. Eystein Husebye
NTNF/NORSAR
P.O. Box 51
N-2007 Kjeller, NORWAY

David Jepsen
Acting Head, Nuclear Monitoring Section
Bureau of Mineral Resources
Geology and Geophysics
G.P.O. Box 378, Canberra, AUSTRALIA

Ms. Eva Johannisson
Senior Research Officer
FOA
S-172 90 Sundbyberg, SWEDEN

Dr. Peter Marshall
Procurement Executive
Ministry of Defense
Blacknest, Brimpton
Reading FG7-FRS, UNITED KINGDOM

Dr. Bernard Massinon, Dr. Pierre Mechler
Societe Radiomana
27 rue Claude Bernard
75005 Paris, FRANCE (2 Copies)